

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: December 30, 2004, 16:50:07 ; Search time 189 Seconds

(without alignments)
54.798 Million cell updates/sec

Title: US-10-031-289-1331

Perfect score: 88

Sequence: 1 PTQAAELNKSKELEQQ 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	100.0	1975	2 Q9K087	Q9K087 neisseria m
2	88	100.0	1995	2 Q9JY23	Q9JY23 neisseria m
3	88	100.0	2015	2 Q7AX69	Q7AX69 neisseria m
4	88	100.0	2015	2 Q9JRD2	Q9JRD2 neisseria m
5	56	53.6	857	2 Q7PML3	Q7PML3 anopheles g
6	52	59.1	59	2 Q03122	Q03122 streptococc
7	52	59.1	163	2 Q9ZFU1	Q9ZFU1 streptococc
8	52	59.1	220	2 Q8GL96	Q8GL96 streptococc
9	51	58.0	506	2 Q8WQ96	Q8WQ96 dictyosteli
10	50	56.8	211	2 Q6YQJ6	Q6YQJ6 onion yello
11	50	56.8	211	2 BAD04462	BAD04462 onion yel
12	49	55.7	246	2 Q3MF63	Q3MF63 beta vulgar
13	49	55.7	264	2 Q6URW3	Q6URW3 streptococc
14	49	55.7	264	2 Q9JRD2	Q9JRD2 streptococc
15	49	55.7	592	2 Q00720	Q00720 streptococc
16	49	55.7	2469	1 TEGU SHV21	T01056 salmireline
17	48	54.5	241	2 Q9TS99	Q9TS99 erinaceus e
18	48	54.5	2477	2 Q80B16	Q80B16 salmireline
19	47	53.4	199	2 Q6FPA3	Q6FPA3 candida gla
20	47	53.4	566	1 SNFC YEAST	SNFC3628 saccharomyc
21	46	52.3	171	2 Q8R759	Q8R759 thermoanaer
22	46	52.3	207	2 Q6YPS4	Q6YPS4 onion yello
23	46	52.3	207	2 Q6YR88	Q6YR88 onion yello
24	46	52.3	207	2 BAD04152	BAD04152 onion yel
25	46	52.3	207	2 BAD04736	BAD04736 onion yel
26	46	52.3	222	2 Q6INL9	Q6INL9 xenopus lae
27	46	52.3	222	2 AAH72260	AAH72260 xenopus l
28	46	52.3	243	2 Q6YQJ3	Q6YQJ3 onion yello
29	46	52.3	243	2 BAD04465	BAD04465 onion yel
30	46	52.3	333	2 Q8YUV3	Q8YUV3 anabaena sp
31	46	52.3	460	2 Q8ENJ2	Q8ENJ2 oceanobacil

32	46	52.3	810	2	Q99207	Q99207 saccharomyc
33	46	52.3	1318	2	Q7UG24	Q7UG24 rhodospirillum
34	45.5	51.7	752	2	Q9NHD9	Q9NHD9 ceratitidis c
35	45	51.1	172	2	Q9WZV4	Q9WZV4 thermotoga
36	45	51.1	215	2	Q6YQMS	Q6YQMS onion yello
37	45	51.1	215	2	BAD04433	BAD04433 onion yel
38	45	51.1	256	2	Q6BFH4	Q6BFH4 paramacium
39	45	51.1	296	2	Q9SNU6	Q9SNU6 oryza sativ
40	45	51.1	317	2	Q8RL22	Q8RL22 providencia
41	45	51.1	352	2	Q871R3	Q871R3 neurospora
42	45	51.1	440	1	BIK1 YEAST	BIK1 YEAST
43	45	51.1	487	2	Q7XSX9	Q7XSX9 oryza sativ
44	45	51.1	589	1	YHP4 SCHPO	YHP4 SCHPO
45	45	51.1	631	2	Q9KSG7	Q9KSG7 vibrio chol
46	45	51.1	668	2	Q97KW1	Q97KW1 clostridium
47	45	51.1	679	2	Q7S7J4	Q7S7J4 caenorhabdi
48	45	51.1	819	2	Q93228	Q93228 caenorhabdi
49	45	51.1	1004	2	Q8MF26	Q8MF26 dictyosteli
50	45	51.1	1030	2	Q7XWB6	Q7XWB6 oryza sativ
51	45	51.1	1225	2	Q8RJI1	Q8RJI1 incj plasmid
52	45	51.1	1831	2	Q7XN10	Q7XN10 oryza sativ
53	45	51.1	2066	2	Q8LTH9	Q8LTH9 staphylococ
54	45	51.1	2066	2	Q8SDP3	Q8SDP3 staphylococ
55	45	51.1	2066	2	Q6GAK2	Q6GAK2 staphylococ
56	45	51.1	2066	2	Q6GGR4	Q6GGR4 staphylococ
57	45	51.1	2066	2	Q8NWK8	Q8NWK8 staphylococ
58	45	51.1	2067	2	Q9B0D2	Q9B0D2 staphylococ
59	45	51.1	2067	2	BAB21743	BAB21743 staphyloc
60	45	51.1	4499	1	DYHA CHLRE	DYHA CHLRE
61	44	50.0	98	2	Q82XR2	Q82XR2 nitrosomon
62	44	50.0	158	2	Q6CWT4	Q6CWT4 kluyveromyc
63	44	50.0	199	1	TPM1 YEAST	TPM1 YEAST
64	44	50.0	220	2	Q84QA2	Q84QA2 oryza sativ
65	44	50.0	304	2	Q9WZD9	Q9WZD9 thermotoga
66	44	50.0	410	1	YIBP HAEIN	YIBP HAEIN
67	44	50.0	451	2	Q8Z6G9	Q8Z6G9 haemophilus
68	44	50.0	451	2	Q8ZPU2	Q8ZPU2 salmonella
69	44	50.0	461	2	Q8YR84	Q8YR84 anabaena sp
70	44	50.0	468	2	Q8FH09	Q8FH09 escherichia
71	44	50.0	495	2	Q9FYG6	Q9FYG6 arabidopsis
72	44	50.0	553	2	Q6CPQ5	Q6CPQ5 kluyveromyc
73	44	50.0	629	1	RA21_XENLA	RA21_XENLA
74	44	50.0	629	2	Q6DCL7	Q6DCL7 xenopus lae
75	44	50.0	926	2	Q22515	Q22515 caenorhabdi
76	44	50.0	991	2	Q6F4C5	Q6F4C5 babesia equ
77	44	50.0	1089	2	Q6DDI6	Q6DDI6 xenopus lae
78	44	50.0	1441	2	Q7RGY2	Q7RGY2 plasmodium
79	44	50.0	1585	1	YQBO_BACSU	YQBO_BACSU
80	43	48.9	75	2	Q54581	Q54581 streptococc
81	43	48.9	143	2	Q85270	Q85270 streptococc
82	43	48.9	152	2	Q04205	Q04205 arabidopsis
83	43	48.9	156	2	Q8Y8Q6	Q8Y8Q6 listeria mo
84	43	48.9	212	2	Q813K1	Q813K1 bacillus ce
85	43	48.9	228	2	Q8GW87	Q8GW87 arabidopsis
86	43	48.9	272	1	YDBA_BACSU	YDBA_BACSU
87	43	48.9	287	2	Q8GLA9	Q8GLA9 streptococc
88	43	48.9	322	1	YQBD_BACSU	YQBD_BACSU
89	43	48.9	323	2	Q9SHJ6	Q9SHJ6 bacillus su
90	43	48.9	372	2	Q31406	Q31406 gallus gall
91	43	48.9	413	2	Q8BUC6	Q8BUC6 mus musculu
92	43	48.9	418	2	Q9D2Y0	Q9D2Y0 mus musculu
93	43	48.9	443	2	Q8YMQ3	Q8YMQ3 anabaena sp
94	43	48.9	488	1	SYK MYCHO	SYK MYCHO
95	43	48.9	496	2	Q6P0V1	Q6P0V1 brachydanio
96	43	48.9	496	2	Q7ZUY4	Q7ZUY4 brachydanio
97	43	48.9	524	2	AAH65437	AAH65437 brachydan
98	43	48.9	524	2	Q8IHE1	Q8IHE1 drosophila
99	43	48.9	524	2	Q9VTE9	Q9VTE9 drosophila
100	43	48.9	560	2	O58015	O58015 pyrococcus

ALIGNMENTS

```

RESULT 1
Q9K057 PRELIMINARY; PRT; 1975 AA.
AC Q9K057;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hemagglutinin/hemolysin-related protein.
GN OrderedLocusNames=NM00497;
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
[1]
SEQUENCE FROM N.A.
RP STRAIN=MC58 / Serogroup B;
RC MEDLINE=20175755; PubMed=10710307;
RX Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA Cittone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizzi M.,
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815 (2000).
DR EMBL; AE002406; AAF40929.1; -.
DR PIR; B81192; B81192.
DR TIGR; NMB0497; -.
DR InterPro; IPR006915; DUF637.
DR InterPro; IPR006914; DUF638.
DR InterPro; IPR008619; Fil haemagg.
DR InterPro; IPR008638; Haemagg act N.
DR InterPro; IPR009013; Viral_att_shaft.
DR Pfam; PF04830; DUF637; 1.
DR Pfam; PF04829; DUF638; 1.
DR Pfam; PF05594; Fil haemagg; 7.
DR Pfam; PF05860; Haemagg act; 1.
DR TIGRFAMs; TIGR01901; adhes_NPXG; 1.
DR Complete proteome.
KW SEQUENCE 1975 AA; 204674 MW; 972CD7BED908EF0B CRC64;
SQ
Query Match 100.0%; Score 88; DB 2; Length 1975;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQ 18
Db 956 PTOKAAELNOKSKELEQ 973

RESULT 2
Q9JY23 PRELIMINARY; PRT; 1995 AA.
AC Q9JY23;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hemagglutinin/hemolysin-related protein.
GN OrderedLocusNames=NM01779;
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
[1]
SEQUENCE FROM N.A.
RP STRAIN=MC58 / Serogroup B;
RC MEDLINE=20175755; PubMed=10710307;
RX Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA Cittone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizzi M.,
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815 (2000).
DR EMBL; AE002406; AAF40929.1; -.
DR PIR; B81192; B81192.
DR TIGR; NMB0497; -.
DR InterPro; IPR006915; DUF637.
DR InterPro; IPR006914; DUF638.
DR InterPro; IPR008619; Fil haemagg.
DR InterPro; IPR008638; Haemagg act N.
DR InterPro; IPR009013; Viral_att_shaft.
DR Pfam; PF04830; DUF637; 1.
DR Pfam; PF04829; DUF638; 1.
DR Pfam; PF05594; Fil haemagg; 7.
DR Pfam; PF05860; Haemagg act; 1.
DR TIGRFAMs; TIGR01901; adhes_NPXG; 1.
DR Complete proteome.
KW SEQUENCE 1975 AA; 204674 MW; 972CD7BED908EF0B CRC64;
SQ
Query Match 100.0%; Score 88; DB 2; Length 1975;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQ 18
Db 956 PTOKAAELNOKSKELEQ 973

```

```

RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA Cittone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizzi M.,
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815 (2000).
DR EMBL; AE002527; AAF42119.1; -.
DR PIR; G81044; G81044.
DR TIGR; NMB1779; -.
DR InterPro; IPR006915; DUF637.
DR InterPro; IPR006914; DUF638.
DR InterPro; IPR008619; Fil haemagg.
DR InterPro; IPR008638; Haemagg act N.
DR InterPro; IPR009013; Viral_att_shaft.
DR Pfam; PF04830; DUF637; 1.
DR Pfam; PF04829; DUF638; 1.
DR Pfam; PF05594; Fil haemagg; 7.
DR Pfam; PF05860; Haemagg act; 1.
DR TIGRFAMs; TIGR01901; adhes_NPXG; 1.
DR Complete proteome.
KW SEQUENCE 1995 AA; 207211 MW; F40F81F28357285F CRC64;
SQ
Query Match 100.0%; Score 88; DB 2; Length 1995;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQ 18
Db 959 PTOKAAELNOKSKELEQ 976

RESULT 3
Q7AX69 PRELIMINARY; PRT; 2015 AA.
AC Q7AX69;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PhA8 protein.
GN Name=PhA8;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
[1]
SEQUENCE FROM N.A.
RP STRAIN=22491;
RC MEDLINE=20187481; PubMed=10722605;
RX Klee S.R., Nassif X., Kusecek B., Merker P., Beretti J.L., Achtman M.,
RA Tinsley C.R.;
RA "Molecular and biological analysis of eight genetic islands that
RT distinguish Neisseria meningitidis from the closely related pathogen
RL Infect. Immun. 68:2082-2095 (2000).
DR EMBL; AJ391255; CAB71945.1; -.
DR InterPro; IPR006915; DUF637.
DR InterPro; IPR006914; DUF638.
DR InterPro; IPR008619; Fil haemagg.
DR InterPro; IPR008638; Haemagg act N.
DR InterPro; IPR011004; Trimer_LpXA_like.
DR Pfam; PF04830; DUF637; 1.
DR Pfam; PF05594; Fil haemagg; 5.
DR Pfam; PF05860; Haemagg act; 1.
DR TIGRFAMs; TIGR01901; adhes_NPXG; 1.
DR Complete proteome.
KW SEQUENCE 2015 AA; 211382 MW; 3500BDC962BC8B0E CRC64;
SQ
Query Match 100.0%; Score 88; DB 2; Length 2015;

```

Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTQKAAELNOKSKELEQ 18
| | | | | | | | | | | | | | | | | | | | | |
Db 959 PTQKAAELNOKSKELEQ 976

RESULT 4

Q9JRD2 ID Q9JRD2 PRELIMINARY; PRT; 2015 AA.
AC Q9JRD2
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein NMA0688.
GN OrderedLocNames=NMA0688;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491";
RL Nature 404:502-506(2000).
DR EMBL; AL162753; CA883974.1; -.
DR PIR; B81989; B81989.
DR InterPro; IPR006915; DUF637.
DR InterPro; IPR006914; DUF638.
DR InterPro; IPR008619; Fil haemagg.
DR InterPro; IPR008638; Haemagg act N.
DR InterPro; IPR009013; Viral_att_shaft.
DR Pfam; PF04830; DUF637; 1.
DR Pfam; PF04829; DUF638; 1.
DR Pfam; PF05594; Fil haemagg; 5.
DR Pfam; PF05860; Haemagg act; 1.
DR TIGRfams; TIGR01901; adhes NPXG; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 2015 AA; 211382 MW; 3500BDC962BC8B0E CRC64;

Query Match 100.0%; Score 88; DB 2; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTQKAAELNOKSKELEQ 18
| | | | | | | | | | | | | | | | | | | | | |
Db 959 PTQKAAELNOKSKELEQ 976

RESULT 5

Q7PML3 ID Q7PML3 PRELIMINARY; PRT; 857 AA.
AC Q7PML3
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000015879 (Fragment).
GN Name=ENSANGP0000013390;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

DR EMBL; AAB01008978; EAA13643.2; -.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR001752; kinesin motor.
DR InterPro; IPR001220; Lectin_legb.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
DR PROSITE; PS00307; LECTIN LEGUME BETA; UNKNOWN 1.
KW ATP-binding; Microtubule; Motor protein.
FT NON_TER 1
FT NON_TER 857
SQ SEQUENCE 857 AA; 95980 MW; 3DF503FA71E35833 CRC64;

Query Match 63.6%; Score 56; DB 2; Length 857;
Best Local Similarity 70.6%; Pred. No. 12;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TQKAAELNOKSKELEQ 18
| | | | | | | | | | | | | | | | | | | | | |
Db 740 TQKAAELNOKSKELEQ 756

RESULT 6

Q03122 ID Q03122 PRELIMINARY; PRT; 59 AA.
AC Q03122
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE M-like protein (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93084815; PubMed=1339461;
RX Relf W.A., Martin D.R., Sriprakash K.S.;
RT "Identification of sequence types among the M-nontypeable group A
streptococci [see comments].";
RL J. Clin. Microbiol. 30:3190-3194(1992).
DR EMBL; L05024; AAA21790.1; -.
FT NON_TER 1
FT NON_TER 59
SQ SEQUENCE 59 AA; 6996 MW; FA7A45ADA1A26857 CRC64;

Query Match 59.1%; Score 52; DB 2; Length 59;
Best Local Similarity 64.7%; Pred. No. 3.1;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTQKAAELNOKSKELEQ 17
| | | | | | | | | | | | | | | | | | | | | |
Db 7 PVKKAELYDKIKELEE 23

RESULT 7

Q9ZFL1 ID Q9ZFL1 PRELIMINARY; PRT; 163 AA.
AC Q9ZFL1
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2346;
RX MEDLINE=97350988; PubMed=9207364;
RA Fiorentino T.R., Beall B., Mshar P., Bessen D.E.;
RT "A genetic-based evaluation of the principal tissue reservoir for
group A streptococci isolated from normally sterile sites.";
J. Infect. Dis. 176:177-182(1997).
RL [2]
RN SEQUENCE FROM N.A.
RP Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RC Beall B.;
RA EMBL; AF098299; AAC79688.1; -;
DR GO; GO:0016020; C-membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 1.
FT NON TER 1 163
FT NON TER 163 163
SQ SEQUENCE 163 AA; 18862 MW; 596FC3E149FFAD7A CRC64;

Query Match 59.1%; Score 52; DB 2; Length 163;
Best Local Similarity 64.7%; Pred. No. 8.9;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTKAAELNOKSKELEQ 17
| : ||||| | |||||
Db 29 PVKAAELYDKILEE 45

RESULT 8
Q8GL96 PRELIMINARY; PRT; 220 AA.
ID Q8GL96;
AC Q8GL96;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dyal-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,
RA McMillan D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY139411; AAN64684.1; -;
DR GO; GO:0016020; C-membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 2.
FT NON TER 1 220
FT NON TER 220 220
SQ SEQUENCE 220 AA; 25396 MW; DA9F34BC072230AA CRC64;

Query Match 59.1%; Score 52; DB 2; Length 220;
Best Local Similarity 64.7%; Pred. No. 12;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTKAAELNOKSKELEQ 17
| : ||||| | |||||
Db 30 PVKAAELYDKILEE 46

RESULT 9
Q8WQ86 PRELIMINARY; PRT; 506 AA.
ID Q8WQ86;
AC Q8WQ86;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Microtubule-associated protein EBI.
GN Name=ebi;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RA Rehberg M., Graef R.;
RT "DeEB1 is a permanent centrosomal resident in Dictyostelium and
required for proper spindle formation.";
Mol. Biol. Cell 12S:309a-309a(2001).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=AX2;
RX MEDLINE=22129383; PubMed=12134070;
RA Rehberg M., Graef R.;
RT "Dictyostelium EBI is a genuine centrosomal component required for
proper spindle formation.";
Mol. Biol. Cell 13:2301-2310(2002).
RL EMBL; AJ426053; CAD19801.1; -;
DR DictyBase; DDB0191160; EBI.
DR GO; GO:0008017; F-microtubule binding; IEA.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR004953; EBI.
DR Pfam; PF00307; CH; 1.
DR Pfam; PF03271; EBI; 1.
DR SMART; SMO0033; CH; 1.
DR PROSITE; PS0021; CH; 1.
SQ SEQUENCE 506 AA; 56983 MW; F20BC182BAAA639C CRC64;

Query Match 58.0%; Score 51; DB 2; Length 506;
Best Local Similarity 55.6%; Pred. No. 40;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PTKAAELNOKSKELEQ 18
| : ||||| | |||||
Db 283 PTELLELEQKRELEQ 300

RESULT 10
Q6YQJ6 PRELIMINARY; PRT; 211 AA.
ID Q6YQJ6;
AC Q6YQJ6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PAM377;
OS Onion yellows phytoplasma.
OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
OC Acholeplasmataceae; Candidatus Phytoplasma.
OX NCBI_TaxID=100379;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OY-M;
RX PubMed=14661021;
RA Oshima K., Kakizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
RA Arashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
RT "Reductive evolution suggested from the complete genome sequence of a
plant-pathogenic phytoplasma.";
Nat. Genet. 36:27-29(2004).
RL EMBL; AP006628; BAD04462.1; -;
DR EMBL; AP006628; BAD04462.1; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 211 AA; 24750 MW; 2A0959A39DB29DBC CRC64;

Query Match 56.8%; Score 50; DB 2; Length 211;
Best Local Similarity 62.5%; Pred. No. 23;

NCBI - T

```
Query Match      54.5%; Score 48; DB 2; Length 241;
Best Local Similarity 68.8%; Pred. No. 53;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

QY 3 QKAAELNQKSKELEQQ 18
||| ||| : |||
Db 131 QKAQELQQKAGELGQQ 146

RESULT 18

Q80BL6	Q80BL6	PRELIMINARY;	PRT; 2477 AA..
AC	Q80BL6;		
DT	01-JUN-2003	(TEMBLrel. 24, Created)	
DT	01-JUN-2003	(TEMBLrel. 24, Last sequence update)	
DT	01-OCT-2003	(TEMBLrel. 25, Last annotation update)	
DE	Hypothetical protein.		
DE	Saimiriine herpesvirus 2.		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Gammaherpesvirinae; Rhadinovirus.		
OX	NCBI TaxID=10381;		

RN	[1]	SEQUENCE FROM N.A.
RP		

RC	STRAIN=C488;	
RX	MEDLINE=22918177; PubMed=14554077;	
RA	Essner A., Thuru M., Wittmann S., Fickenscher H.;	
RA	"the genome of herpesvirus saimiri C488 which is capable of	
RT	transforming human T cells.;"	
RI		
RL	Virology 314:471-487(2003).	
DR	EMBL; AJ410493; CAC84361.1;	
DR	InterPro; IPR006928; Herpes_teg_N.	
DR	Pfam; PF04843; Herpes_teg_N; 1.	
DR	Hypothetical protein.	
SK	SEQUENCE 2477 AA; 281167 MW; 73AF050823D27639 CRC64;	
QW		

Query Match 54.5%; Score 48; DB 2; Length 2477;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 9: Conservative 6; Mismatches 3; Indels

Qy 1 PTQAAELNQKSKELEQQ 18
|::||:|:||||:|:
nb 671 PSKKAQOOLEKSKLVEOK 688

RESULT 19
06FPA3

ID	Q6FPA3	PRELIMINARY;	PRT;	199 AA.
AC	Q6FPA3;			
DT	05-JUL-2004 (TtEMBLrel. 27, Created)			
DT	05-JUL-2004 (TtEMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TtEMBLrel. 27, Last annotation update)			
DE	Strain CBS138 chromosome J complete sequence.			
GN	ORFNames=CAGJ0J05434g;			
OS	Candida glabrata (Yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; mitosporic Saccharomycetales; Candida.			
OX	NCBI_TaxID=5478;			
XX	[]			
RP	SEQUENCE FROM N.A.			
RN	STRAIN=CBS138;			
RG	GENOLEVURES;			
RA	Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,			
RA	Lafontaine I., de Montigny J., Marcq C., Neuveglise C., Talla E., V.,			
RA	Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,			
RA	Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykassen C.,			
RA	Bolsrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,			
RA	Despons L., Fabre E., Fairhead C., Perry-Dumazet H., Groppi A.,			
RA	Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,			
RA	Karrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,			
RA	Nicaud J.M., Nikolaïski M., Ozias S., Ozier-Kalogeropoulos O.,			
RA	Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,			
RA	Swennene D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,			
RA	Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,			

RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.
RA
RA Winkler P., Souciet J.L.,
RT "genome evolution in yeasts";
RL Nature 430:35-44 (2004)).
RL EMBL; CAG608956; CAG60892.1; -.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYSIN.
SO SEQUENCE 199 AA; 23276 MW; F24AD45FC382477 CRC64;

Query Match 53.4%; Score 47; DB 2; Length 199;
Best Local Similarity 56.2%; Pred. No. 61;
Matches 9; Conservative 4; Mismatches 3; Indels

Qy 3 QKAAELNQSKLEQQ 18
:|:||:|:|:|:|:
Db 20 EKYEELKEKNKELEOE 35

RESULT 20

SNFC_YEAST

ID - SNFC_YEAST STANDARD; PRT; 566 AA.

AC P53628;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Transcription regulatory protein SNF12 (SWI/SNF complex component

DE SNF73)

GN Names=SNF12; Synonyms=SWP73; OrderedLocusNames=YNR023W; ORFNames=N3224

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RN SEQUENCE FROM N.A.

RP

RA Donald K.A.G., Hill J., Griffiths D.E.;

RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.

RN [2]

RN SEQUENCE FROM N.A.

RP

RA Pohl T.M.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases

DD 6/I PSAAUQUUGENSKUVEQA 000

RESULT 19

Genes Dev. 10:2131-2144(1996).

-1- FUNCTION: Involved in transcriptional activation. The SWI/SNF complex is required for the induced expression of a large number of genes. This complex alters chromatin structure to facilitate binding of gene-specific dedicated transcription factors.

-1- SUBUNIT: Component of the SWI/SNF global transcription activator complex.

-1- SUBCELLULAR LOCATION: Nuclear.

-1- SIMILARITY: To yeast RSC6.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/annualreport.html> or send an email to license@isb-sib.ch).

CC	EMBL; X62430; CAA44296.1; --
DR	EMBL; Z71638; CAA96302.1; --
DR	PIR; S19063; S13063.
DR	Interact; P53628; --
DR	GeneOnline; 143368; --
DR	TRANSFAC; T03614; --
DR	SGD; S0005308; SNF12.
DR	Activator; Nuclear protein;
DR	SEQUENCE 556 AA; 63870 MW

DK SGD; 50003306; SMF12.
KW Activator; Nuclear protein; Transcription regulation.
SQ SEQUENCE 566 AA; 63870 MW; CD6E61EEEE85098C CRC64;

Query Match 53.4%; Score 47; DB 1; Length 566;
 Best Local Similarity 69.2%; Pred. No. 1.8e+02;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AAEINQKSKLEEQ 17

Db 463 AAEINENARELEEQ 475

RESULT 21

Q8R7R9 ID Q8R7R9 PRELIMINARY; PRT; 171 AA.
 AC Q8R7R9
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=TFE2329;
 OS Thermoanaerobacter tengcongensis.
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
 OC Thermoanaerobacteraceae; Thermoanaerobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 Chen Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of the T. tengcongensis genome.";
 RL Genome Res. 12:689-700(2002).
 DR EMBL; AE013175; RA025470.1; --
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004518; F:nuclease activity; IEA.
 DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
 DR InterPro; IPR001943; UvrB/C.
 DR InterPro; IPR009055; UvrB_C.
 DR Pfam; PF02151; Uvr; 1.
 DR PROSITE; PS50151; UVR; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 171 AA; 19618 MW; A81865FB9FE41B2D CRC64;

Query Match 52.3%; Score 46; DB 2; Length 171;
 Best Local Similarity 56.2%; Pred. No. 74;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 QKAAELNQKSKLEEQ 18

Db 153 EKAAELRDKIRELEKE 168

RESULT 22

Q6VPS4 ID Q6VPS4 PRELIMINARY; PRT; 207 AA.
 AC Q6VPS4
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=PAM651;
 OS Onion yellows phytoplasma.
 OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
 OC Acholeplasmataceae; Candidatus Phytoplasma.
 OX NCBI_TaxID=100379;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OY-M;
 RX PubMed=14661021;
 RA Oshima K., Kakizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
 Arashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
 RT "Reductive evolution suggested from the complete genome sequence of a
 plant-pathogenic phytoplasma.";

RL Nat. Genet. 36:27-29(2004).
 DR EMBL; AP006628; BAD04736.1; --
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 207 AA; 24543 MW; B9EC9177465AF288 CRC64;

Query Match 52.3%; Score 46; DB 2; Length 207;
 Best Local Similarity 56.2%; Pred. No. 90;
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNQKSKLEEQ 18

Db 75 EQANEINRLSEELEEQ 90

RESULT 23

Q6YRE8 ID Q6YRE8 PRELIMINARY; PRT; 207 AA.
 AC Q6YRE8
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=PAM067;
 OS Onion yellows phytoplasma.
 OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
 OC Acholeplasmataceae; Candidatus Phytoplasma.
 OX NCBI_TaxID=100379;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OY-M;
 RX PubMed=14661021;
 RA Oshima K., Kakizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
 Arashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
 RT "Reductive evolution suggested from the complete genome sequence of a
 plant-pathogenic phytoplasma.";
 RL Nat. Genet. 36:27-29(2004).
 DR EMBL; AP006628; BAD04152.1; --
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 207 AA; 24577 MW; E3469BD7EC505888 CRC64;

Query Match 52.3%; Score 46; DB 2; Length 207;
 Best Local Similarity 56.2%; Pred. No. 90;
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNQKSKLEEQ 18

Db 75 EQANEINRLSEELEEQ 90

RESULT 24

BAD04152 ID BAD04152 PRELIMINARY; PRT; 207 AA.
 AC BAD04152;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Conserved hypothetical protein.
 GN PAM067.
 OS Onion yellows phytoplasma.
 OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
 OC Acholeplasmataceae; Phytoplasma.
 OX NCBI_TaxID=100379;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OY-M;
 RX PubMed=14661021;
 RA Oshima K., Kakizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
 Arashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
 RT "Reductive evolution suggested from the complete genome sequence of a
 plant-pathogenic phytoplasma.";
 RL Nat. Genet. 36:27-29(2004).
 DR EMBL; AP006628; BAD04152.1; --
 KW Hypothetical protein.

Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyszewski M.I., Skalka U., Smalls D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2].
SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
SQ EMBL; BC072260; AAH72260.1; -.
SR SEQUENCE 222 AA; 25110 MW; 2A0C9719F668B5D1 CRC64;
Query Match 52.3%; Score 46; DB 2; Length 222;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 TOKAAELNOKSKELE 16
: : : : :
DB 53 TEIKABINQKSELE 67
RESULT 27
AAH72260
ID AAH72260 PRELIMINARY; PRT; 222 AA.
AC AAH72260;
DT 01-JUN-2004 (T-EMBLrel. 27, Created)
DT 01-JUN-2004 (T-EMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
NCBI_TaxID=8355;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heif F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavani T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.B., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalski U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (3)
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RC Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBDJ databases.
DR EMBL; BC072260; AAH72260.1; -.
KW Hypothetical protein.
SQ SEQUENCE 222 AA; 25110 MW; 2A0C9719F668B5D1 CRC64;

Query Match 52.3%; Score 46; DB 2; Length 222;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TQKAAELNOKSKELE 16
Db 53 TEIKAEINOKSLELE 67

RESULT 28
Q6YQJ3 PRELIMINARY; PRT; 243 AA.
AC Q6YQJ3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Chromosome segregation ATPase homolog.
GN Name=smc; OrderedLocusNames=PAM380;
OS Onion yellows phytoplasma.
OC Bacteria; Firmicutes; Mollicutes; Achaeplasmatales;
OC Achaeplasmataceae; Candidatus Phytoplasma.
OX NCBI_TaxID=100379;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=OY-M;
RX PubMed=14661021;
RA Oshima K., Kakizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
RA Arashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
RT "Reductive evolution suggested from the complete genome sequence of a
RT plant-pathogenic phytoplasma.";
RL Nat. Genet. 36:27-29(2004).
DR EMBL; AP006628; BAD04465.1; -.
KW Complete proteome.
SQ SEQUENCE 243 AA; 28791 MW; C3C3CF555AC5B441 CRC64;

Query Match 52.3%; Score 46; DB 2; Length 243;
Best Local Similarity 56.2%; Pred. No. 1.1e+02;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELE 16
Db 170 EQANEINRLSELEEQ 185

RESULT 29
BAD04465 PRELIMINARY; PRT; 243 AA.
AC BAD04465;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Chromosome segregation ATPase homolog.
GN SMC OR PAM380.
OS Onion yellows phytoplasma.
```

```
OC Bacteria; Firmicutes; Mollicutes; Achaeplasmatales;
OC Achaeplasmataceae; Phytoplasma.
OX NCBI_TaxID=100379;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=OY-M;
RX PubMed=14661021;
RA Oshima K., Kakizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
RA Arashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
RT "Reductive evolution suggested from the complete genome sequence of a
RT plant-pathogenic phytoplasma.";
RL Nat. Genet. 36:27-29(2004).
DR EMBL; AP006628; BAD04465.1; -.
SQ SEQUENCE 243 AA; 28791 MW; C3C3CF555AC5B441 CRC64;

Query Match 52.3%; Score 46; DB 2; Length 243;
Best Local Similarity 56.2%; Pred. No. 1.1e+02;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELEEQ 18
Db 170 EQANEINRLSELEEQ 185

RESULT 30
Q8YUV3 PRELIMINARY; PRT; 333 AA.
AC Q8YUV3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ABC transporter, phosphate-binding protein.
GN Name=phnd; OrderedLocusNames=all2228;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003588; BAB73927.1; -.
DR PIR; AF2084; AF2084.
KW Complete proteome.
SQ SEQUENCE 333 AA; 37146 MW; 1EFPB27FECSA87DC CRC64;

Query Match 52.3%; Score 46; DB 2; Length 333;
Best Local Similarity 64.3%; Pred. No. 1.5e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELE 16
Db 320 QKIAELNQLKEIQ 333

Search completed: December 30, 2004, 17:00:07
Job time : 195 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:46:31 ; Search time 151 Seconds
(without alignments)
42.762 Million cell updates/sec

Title: US-10-031-289-1331

Perfect score: 88

Sequence: 1 PTKAAELNKSKELEQQ 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	1978	2	AA27230 Amino aci
2	88	100.0	1981	2	AAW42634 Protein s
3	88	100.0	2015	5	ABB78067 Amino aci
4	88	100.0	2015	6	ABU37848 Protein e
5	67	76.1	1532	2	AA27231 Amino aci
6	48	54.5	180	8	ADJ93927 Western E
7	47	53.4	139	5	ABP10237 Human ORF
8	47	53.4	566	7	ABR53129 Protein s
9	47	53.4	566	7	ADK62560 Disease t
10	46	52.3	350	4	AAE00658 Chicken p
11	46	52.3	810	5	AAE24680 Yeast Emg
12	46	52.3	810	5	ABR53226 Protein s
13	46	52.3	810	7	ADK63710 Disease t
14	45	51.1	285	4	AA663172 Amino aci
15	45	51.1	440	1	AAAP94878 BIK 1 inc
16	45	51.1	440	2	AAAR14907 BIK1 prot
17	45	51.1	1905	8	ADI45155 Rice isop
18	45	51.1	2074	6	ABM70790 Staphyloc
19	44	50.0	43	2	AAW69983 Yeast tro
20	44	50.0	410	5	AAO17657 H influen
21	44	50.0	410	5	AAO17658 H influen
22	44	50.0	410	5	AAO17659 H influen
23	44	50.0	410	5	AAO17656 H influen
24	44	50.0	1224	6	ABDI0047 Alloiococ
25	43	48.9	94	2	AA21561 dnaG carb

26	43	48.9	152	3	AAG30644	Aag30644 Arabidops
27	43	48.9	156	5	ABB47539	Abb47539 Listeria
28	43	48.9	219	3	AAG30643	Aag30643 Arabidops
29	43	48.9	228	3	AAG30642	Aag30642 Arabidops
30	43	48.9	323	3	AAG41490	Aag41490 Arabidops
31	43	48.9	524	4	ABB63524	Abb63524 Drosophil
32	43	48.9	567	6	ADB08174	Adb08174 Alloiococ
33	43	48.9	574	6	ADB08176	Adb08176 Alloiococ
34	43	48.9	626	5	ABU49107	Abu49107 Listeria
35	43	48.9	626	6	ABU32651	Abu32651 Protein e
36	43	48.9	1013	6	ABU49691	Abu49691 Protein e
37	43	48.9	2665	4	AAAI4533	Aami4533 Peptide #
38	43	48.9	2665	4	ABB33490	Abb33490 Peptide #
39	43	48.9	2665	4	AAAM26950	Aam26950 Peptide #
40	43	48.9	2665	4	ABB28314	Abb28314 Human pep
41	43	48.9	2665	4	ABB18950	Abb18950 Protein #
42	43	48.9	2665	4	AAAM6665	Aam6665 Human bon
43	43	48.9	2665	4	AAAM54270	Aam54270 Human bra
44	43	48.9	2665	4	ABAG48336	Abg48336 Human liv
45	43	48.9	2665	4	AAAM02259	Aam02259 Peptide #
46	43	48.9	2665	5	ABG36319	Abg36319 Human pep
47	43	48.9	3266	3	AAAB42491	Aab42491 Human ORF
48	43	48.9	3271	6	ABO07211	Ab007211 Human p53
49	43	48.9	3664	6	ABO47592	Ab047592 Breast ca
50	43	48.9	3664	6	ABO53027	Ab053027 Human put
51	43	48.9	3664	7	ADJ70481	Adj70481 Human hea
52	43	48.9	3664	8	ADK60211	Adk60211 Angiogene
53	43	48.9	3664	8	ADK60512	Adk60512 Angiogene
54	43	48.9	3664	8	ADO01005	Ado01005 Human hom
55	43	48.9	3664	8	ADP73135	Adp73135 Angiogene
56	42.5	48.3	119	7	ADD13010	Add13010 A. gossyp
57	42.5	48.3	152	7	ADD13013	Add13013 A. gossyp
58	42	47.7	25	3	AAAB08381	Aab08381 Peptide u
59	42	47.7	35	3	AAAB08365	Aab08365 Amino aci
60	42	47.7	35	3	AAAB08367	Aab08367 Amino aci
61	42	47.7	35	3	AAAB08345	Aab08345 Amino aci
62	42	47.7	35	3	AAAB08363	Aab08363 Amino aci
63	42	47.7	35	3	AAAB08384	Aab08384 Peptide u
64	42	47.7	35	6	ADA00697	Ada00697 Protein m
65	42	47.7	35	6	ADA00670	Ada00670 Protein m
66	42	47.7	35	6	ADA00691	Ada00691 Protein m
67	42	47.7	35	6	ADA00693	Ada00693 Protein m
68	42	47.7	35	6	ADA00672	Ada00672 Protein m
69	42	47.7	35	6	ADA00653	Ada00653 Protein m
70	42	47.7	35	6	ADA00695	Ada00695 Protein m
71	42	47.7	35	6	ADA00686	Ada00686 Protein m
72	42	47.7	35	6	ADA00687	Ada00687 Protein m
73	42	47.7	35	6	ADA00674	Ada00674 Protein m
74	42	47.7	43	3	AAAB08383	Aab08383 Peptide u
75	42	47.7	43	3	AAAB08364	Aab08364 Amino aci
76	42	47.7	43	6	ADA00694	Ada00694 Protein m
77	42	47.7	43	6	ADA00680	Ada00680 Protein m
78	42	47.7	43	6	ADA00671	Ada00671 Protein m
79	42	47.7	108	3	AAAG01833	Aag01833 Human sec
80	42	47.7	125	8	ADL83152	Adl83152 Human PRO
81	42	47.7	137	4	ABB70287	Abb70287 Drosophil
82	42	47.7	233	4	AAAB95513	Aab95513 Human rep
83	42	47.7	233	4	ABB96196	Abb96196 Human tes
84	42	47.7	287	7	ADM05004	Adm05004 Human pro
85	42	47.7	306	5	ABP41844	Abp41844 Human ova
86	42	47.7	314	8	ADI42800	Adi42800 Plant tra
87	42	47.7	315	6	ABR40787	AbR40787 Brassica
88	42	47.7	321	4	AAU30277	Aau30277 Novel hum
89	42	47.7	327	3	AAAB56631	Aab56631 Human pro
90	42	47.7	327	3	ABU16682	Abu16682 Protein e
91	42	47.7	329	2	AAAY30522	Aay30522 A human m
92	42	47.7	329	4	AAAB62699	Aab62699 Human mem
93	42	47.7	329	5	AAAB71409	Aab71409 Human HMR
94	42	47.7	329	8	ABR59719	AbR59719 Human sec
95	42	47.7	329	8	ADO43996	Ado43996 Amino aci
96	42	47.7	331	7	ADC31639	Adc31639 Human nov
97	42	47.7	347	6	ABU22017	Abu22017 Protein e
98	42	47.7	359	6	ADA35810	Ada35810 Acinetoba

99	42	47.7	383	2	AAW19591	Papaya ri
100	42	47.7	434	7	ADE08569	Novel pro

ALIGNMENTS

RESULT 1
AAY27230
ID AAY27230 standard: protein: 1978 AA.

XX	AAV27230;
AC	XX
XX	24-SEP-1999 (first entry)
DT	XX
XX	Amino acid sequence of N. meningitidis protein ORF114-1.
DE	XX
XX	Neisseria meningitidis protein; pharmaceutical; vaccine; diagnosis;
KW	XX
KW	bacterial infection; treatment.
XX	XX
OS	Neisseria meningitidis.

```

XX
SQ Sequence 1978 AA;

Query Match          100.0%; Score 88; DB 2; Length 1978;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 2
AAW42634
ID AAW42634 standard; protein; 1981 AA.
XX
AC AAW42634;

XX 22-OCT-1998 (first entry)
DT
XX Protein sequence that is specific for Neisseria meningitidis.
DE

XX	N. gonorrhoeae; N. lactamica; chromosome Z2491; region 1; region 2;
KW	region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
KW	meningitis.
XX	
XX	Neisseria meningitidis.
OS	
XX	
XX	WO9802547-A2.
PN	
XX	
XX	22-JAN-1998.
PD	
XX	
XX	11-JUL-1997; 97WO-FR001295.
PF	
XX	
XX	12-JUL-1996; 96FR-00008768.
PR	
XX	
XX	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA	(ELAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA	(SMIK) SMITHKLINE BEECHAM.
PA	
XX	
XX	Nassif X, Tinsley C, Achtman M, Ruelle J, Vinals C, Merker P;
PI	
XX	
XX	WPI; 1998-110594/10.
DR	
XX	N-PSDB: AAV03553.
DR	

Genes present in *Neisseria meningitidis* but not other *Neisseria* species - and related host cells, RNA, anti-sense sequences, polypeptide(s) and antibodies, useful for diagnosing *Neisseria meningitidis* infection and in protective vaccines.

Claim 8: Page 81-90: 150pp: French:

AAW42633-37 and AAW42639-41 are encoded by a DNA sequence found in region 2 of *Neisseria meningitidis*. The specification describes DNA sequences that are found in *N. meningitidis*, but not in *N. gonorrhoeae* or *N. lactamica*, except for the genes involved in biosynthesis of the capsule polysaccharide, fpa or C, opa, rotamase, sequence IC1106, Iga protease, pilin, pilC, proteins which bind transferrin and opacity proteins. The DNA sequences are found on chromosome 22491, mainly (or within 20 kb) between tufa and pilt (region 1), pilQ and lambda-740 (region 2) or argF and opaB (region 3). The DNA sequences are responsible for the differences in pathogenicity between *N. meningitidis* and *N. gonorrhoeae*, specifically they include the genes that allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N. lactamica*, are responsible for colonisation and penetration of the mucosa. The DNA sequences can be used to produce probes and primers, and antibodies produced against the encoded proteins are used in standard hybridisation/immunoassay processes for diagnosis of *N. meningitidis* infection, particularly meningitis.

Sequence 1981 AA:

```

Query Match      100.0%; Score 88; DB 2; Length 1981;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 3	
ABB78067	
ID	ABB78067 standard; protein; 2015 AA.
XX	
XX	
AC	ABB78067;
XX	
DT	29-AUG-2003 (revised)
DT	05-NOV-2002 (first entry)

```
XX OS Neisseria gonorrhoeae.
XX PN WO200260936-A2.
XX PD 08-AUG-2002.
XX PF 31-JAN-2002; 2002WO-US002881.
XX PR 31-JAN-2001; 2001US-0266070P.
XX PR 06-AUG-2001; 2001US-0310356P.
XX PR 23-OCT-2001; 2001US-0344452P.
XX PA (IOWA ) UNIV IOWA RES FOUND.
XX PA (REGC ) UNIV CALIFORNIA.
XX PA (APIC/) APICELLA M A.
XX PA (EDWA/) EDWARDS J L.
XX PA (GIBS/) GIBSON B W.
XX PA (SCHE/) SCHEFFLER K.
XX PA (BROW/) BROWN E.
XX PI Apicella MA, Edwards JL, Gibson BW, Scheffler K, Brown E;
XX DR WPI; 2002-619227/66.
XX DR N-PSDB; ABQ78298.
XX PR New polypeptide comprising p177, p88, p64, p55 or p46 from Neisseria
XX PT gonorrhoeae, useful for preventing, or protecting a female patient
XX PT against, N. gonorrhoeae colonization or infection.
XX FS Claim 7; Page 108-115; 130pp; English.
XX CC The present sequence represents a p177 polypeptide. The specification
XX CC describes p177, p88, p64, p55 and p46 polypeptides from Neisseria
XX CC gonorrhoeae. The polypeptides are useful as vaccines, for preventing, or
XX CC protecting a female patient against, N. gonorrhoeae colonization or
XX CC infection. Such immunisation can prevent gonorrhoea in women. (Updated on
XX CC 29-AUG-2003 to standardise OS field)
XX SQ Sequence 2015 AA;
Query Match 100.0%; Score 88; DB 5; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PTOKAAELNOKSKELEQQ 18
Db 959 PTOKAAELNOKSKELEQQ 976
RESULT 4
ABU37848
ID ABU37848 standard; protein; 2015 AA.
XX AC ABU37848;
XX 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #23375.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Neisseria meningitidis.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00915242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
PA (ELIT-) ELITRA PHARM INC.
XX FI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA41718.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 65772; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 2015 AA;
Query Match 100.0%; Score 88; DB 6; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PTOKAAELNOKSKELEQQ 18
Db 959 PTOKAAELNOKSKELEQQ 976
RESULT 5
AAAY27231
ID AAAY27231 standard; protein; 1532 AA.
XX AC AAAY27231;
XX 24-SEP-1999 (first entry)
XX DE Amino acid sequence of N. meningitidis protein ORF114a.
XX KW Neisseria meningitidis protein; pharmaceutical; vaccine; diagnosis;
XX KW bacterial infection; treatment.
```

OS Neisseria meningitidis.
 XX WO9936544-A2.
 XX 22-JUL-1999.
 XX 14-JAN-1999; 99WO-IB000103.
 XX 14-JAN-1998; 98GB-00000760.
 PR 01-SEP-1998; 98GB-00019015.
 PR 09-OCT-1998; 98GB-00022143.
 XX (CHIR-) CHIRON SPA.
 XX
 PI Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
 XX WPI; 1999-444400/37.
 DR N-PSDB; AAX99153.
 XX
 XX New protein and its nucleotide sequence, useful in vaccines or diagnostic
 PT compositions for treating and/or preventing Neisseria meningitidis
 PT infections.
 XX
 XX Claim 4; Page 91; 123pp; English.
 PS
 CC The invention provides proteins (AAV27201-245) from Neisseria
 CC meningitidis (strains A and B) and nucleic acid sequences (AAX99123-167)
 CC encoding the proteins. Compositions comprising the protein, nucleic acid
 CC or antibody specific to the protein are useful as pharmaceuticals, e.g. a
 CC vaccine composition or a diagnostic composition. The composition is also
 CC useful for treating or preventing an infection due to Neisseria
 CC bacteria, especially Neisseria meningitidis
 CC
 XX Sequence 1532 AA;
 SQ
 Query Match 76.1%; Score 67; DB 2; Length 1532;
 Best Local Similarity 82.4%; Pred. No. 0.28;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 TQKAAELNOKSKELEQQ 18
 Db 960 TQKXXLNOKSKELEQQ 976
 |||||
 RESULT 6
 ADJ93927
 ID ADJ93927 standard; protein; 180 AA.
 XX AC ADJ93927;
 XX
 XX 06-MAY-2004 (first entry)
 XX
 XX Western European hedgehog apolipoprotein (apo-I) protein.
 XX
 XX T-cell protein; immune activation; xenoantigen; immunosuppressive;
 KW antiallergic; antimicrobial; cytostatic; antirheumatic; antiarthritic;
 KW dermatological; neuroprotective; antiinflammatory; vasotropic;
 KW antidabetic; gastrointestinal; antichyroid; antisthmatic; vulnary;
 KW Th1; Th2 immune response; organ transplant rejection; bone marrow;
 KW rheumatoid arthritis; lupus erythematosus; multiple sclerosis;
 KW encephalitis; vasculitis; diabetes mellitus; pancreatitis; gastritis;
 KW thyroiditis; B-cell; asthma; lepromatosis;
 KW Helicobacter pylori-associated gastritis; skin; adrenal; lung tumour;
 KW wound healing; growth disorder; inflammatory; infectious disease;
 KW gene therapy; leucocyte; lymphocyte activation; human;
 KW western European hedgehog; apolipoprotein; apo-I.
 XX
 XX Erinaceus europaeus.
 OS
 XX
 XX WO2004011496-A2.
 XX
 XX 05-FEB-2004.
 PD
 XX

PF 04-JUL-2002; 2002WO-EP007440.
 XX
 PR 04-JUL-2002; 2002WO-EP007440.
 XX
 XX (UTKU/) UTKU N.
 XX
 XX Utku N;
 PI
 XX WPI; 2004-143821/14.
 DR
 XX New polynucleotides encoding T-cell proteins, useful for diagnosing or
 PT treating autoimmune, allergic or infectious diseases or tumors, or for
 PT improving allograft or xenograft tolerance.
 PT
 XX Disclosure; SEQ ID NO 38; 128pp; English.
 PS
 XX The invention relates to a novel polynucleotide encoding a T-cell protein
 CC the expression of which is upregulated during the early stages of immune
 CC activation in response to xenoantigens or their biologically active
 CC fragments. The polynucleotide of the invention demonstrates
 CC immunosuppressive, antiallergic, antidiabetic, cytostatic,
 CC antirheumatic, antiarthritic, dermatological, neuroprotective,
 CC antinflammatory, vasotropic, antidiabetic, gastrointestinal,
 CC antithyroid, antiasthmatic and vulnary activities. The nucleic acid may
 CC be useful for the preparation of a composition for diagnosing or treating
 CC acute and chronic diseases involving T-cell activation and Th1 and Th2
 CC immune response, acute and chronic rejection of allo- and xeno-organ
 CC transplants and bone marrow transplantation, rheumatoid arthritis, lupus
 CC erythematosus, multiple sclerosis, encephalitis, vasculitis, diabetes
 CC mellitus, pancreatitis, gastritis, thyroiditis, malignant disorders of T,
 CC B or NK cells, asthma, lepromatosis, Helicobacter pylori-associated
 CC gastritis, skin, adrenal or lung tumours, in wound healing, growth
 CC disorders and inflammatory and/or infectious diseases, as well as during
 CC gene therapy. The polynucleotide or antibody may also be used for
 CC detecting leucocyte or lymphocyte activation. The current sequence is
 CC that of the western European hedgehog apolipoprotein (apo-I) protein of
 CC the invention.
 XX
 XX Sequence 180 AA;
 SQ
 Query Match 54.5%; Score 48; DB 8; Length 180;
 Best Local Similarity 68.8%; Pred. No. 23;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Oy 3 QKAAELNOKSKELEQQ 18
 Db 131 QKAAELNOKSKELEQQ 146
 |||||
 RESULT 7
 ABP10237
 ID ABP10237 standard; protein; 139 AA.
 XX AC ABP10237;
 XX
 XX 24-JUN-2002 (first entry)
 XX
 XX Human ORFX protein sequence SEQ ID NO:20456.
 XX
 XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypercension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200192523-A2.
 XX
 XX 06-DEC-2001.
 PD

XX New proteins and protein complexes from eukaryotes, useful as targets in
 PT drug screening, or in diagnosing or screening for the presence of a
 PT disease or disorder, or a predisposition for developing a disease or
 PT disorder in a subject.
 XX
 PS Disclosure; SEQ ID NO 811; 13pp; English.
 XX
 CC The invention relates to novel protein complexes comprising a first and a
 CC second protein, or its derivative, fragment, homologue or variant. The
 CC proteins are selected from given protein complexes, which are not defined
 CC in the specification. The variants are encoded by nucleic acids that
 CC hybridize to the nucleic acids encoding the proteins under low stringency
 CC conditions. The protein complexes are useful as targets for an active
 CC agent of a pharmaceutical. These protein complexes are particularly
 CC useful as drug targets for the treatment or preventing of a disease or
 CC disorder. The complexes and methods above are useful in diagnosing or
 CC screening for the presence of a disease or disorder or a predisposition
 CC for developing a disease or disorder in a subject. These are also useful
 CC in screening for a drug for treatment or prevention of a disease or
 CC disorder. The molecule that modulates the amount, activity or protein
 CC components of the complex is useful for the manufacture of a medicament
 CC for the treatment or prevention of a disease or disorder. This sequence
 CC corresponds to a protein of the invention. (Note: the sequence data for
 CC this patent did not form part of the printed specification but was
 CC obtained from the EPO in electronic format).
 XX
 SQ Sequence 566 AA;

Query Match 53.4%; Score 47; DB 7; Length 566;
 Best Local Similarity 69.2%; Pred. No. 1.2e+02;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AAELNQKSKELEQ 17
 |||||: :|||
 Db 463 AAELNENARELEQ 475

RESULT 10
 AAEE00658
 ID AAEE00658 standard; protein; 350 AA.

AC AAEE00658;

DT 02-JUL-2001 (first entry)

DE Chicken protein from clone g16.

XX Chicken; haplotype; B-G antigen; major histocompatibility complex; MHC;
 KW domesticated fowl; restriction fragment length polymorphism; RFLP;
 KW genotyping; clone g16; B system; Marek's disease; immunity.

OS Gallus sp.

XX US6218106-B1.

PN 17-APR-2001.

XX 18-JAN-1995; 95US-00374219.

XX 30-JUN-1987; 87US-00068176.

PR 09-DEC-1987; 87US-00130529.

PR 23-JUN-1988; 88US-00210405.

PR 28-SEP-1989; 89US-00413301.

PR 27-SEP-1990; 90US-00588922.

PR 22-APR-1991; 91US-00688326.

PR 07-APR-1992; 92US-00865662.

XX (CITY) CITY OF HOPE.

FA Miller MM;

XX WPI; 2001-289826/30.

DR N-PSDB; AAD03981.

XX Genotyping or determining the haplotype of domesticated fowls, involves
 PT hybridizing a probe complementary to the B-G sub-region of the major
 PT histocompatibility complex of the fowl to cleavage reaction products.

XX Disclosure; Fig 17; 61pp; English.

XX The present invention relates to a method for determining the haplotype
 CC of a fowl of the order Galliformes or Anseriformes. The method comprises
 CC providing a DNA sample from the fowl, cleaving the sample with a
 CC restriction enzyme, fractionating the products, subjecting the products
 CC to electrophoresis on a gel and hybridising a probe complementary to the
 CC B-G sub-region of the major histocompatibility complex (MHC) of the fowl
 CC to the products. The method is useful for genotyping domesticated fowls
 CC such as chicken, turkey or pheasant for the MHC B-G loci. In this method,
 CC interpretation of results is generally simpler and more uniform since
 CC typing by restriction fragment length polymorphism patterns (RFLP) is no
 CC longer dependent upon alloantisera which often require selective
 CC absorptions with blood samples from genetically-defined animals to
 CC delineate haplotype specificity. The present sequence is a protein
 CC encoded by g16 cDNA clone from chicken of haplotype 21. The cDNA is used
 CC as a probe in the method of the invention. The chicken MHC known as B
 CC system contains three subregions, B-L, B-F and B-G. B-G polypeptide is
 CC used to impart immunity to or to control the susceptibility of
 CC domesticated fowl to various diseases. Resistance to Marek's disease is
 CC closely related to the domesticated fowl MHC

XX Sequence 350 AA;

Query Match 52.3%; Score 46; DB 4; Length 350;

Best Local Similarity 64.3%; Pred. No. 98;

Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AAELNQKSKELEQ 18
 |||||: :|||
 Db 257 AAKLGQQTKELEQ 270

RESULT 11

AAE24680

ID AAE24680 standard; protein; 810 AA.

AC AAE24680;

DT 22-OCT-2002 (first entry)

DE Yeast Emg1-nuclear interacting protein-1 (ENIPI1).

XX Essential for mitotic growth 1; cell cycle regulatory gene; therapy;
 KW Emg1-nuclear interacting protein-1; ENIPI1; neurological disease; Emg1;
 KW Alzheimer's disease; Parkinson's disease; liver cirrhosis; herbicide;
 KW tissue degeneration; fungal infection; cancer; Emg1-1; Yeast.

OS Saccharomyces cerevisiae.

PN US6383753-B1.

XX 07-MAY-2002.

XX 31-MAR-2000; 2000US-00540824.

PR 31-MAR-1999; 99US-0127226P.

XX (UNMI) UNIV MICHIGAN.

PI Thiele DJ, Liu PCC;

XX WPI; 2002-498474/53.

DR N-PSDB; AAD39763.

XX Composition used for identifying potential pharmaceuticals e.g.
 PT anticancer agents comprises isolated and purified DNA having specific

Query Match 52.3%; Score 46; DB 7; Length 810;
Best Local Similarity 56.2%; Pred. No. 2.5e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 QKAAELNOKSKELEQQ 18
DB 294 EKNAEAEKKRELEQQ 309

RESULT 14
AAG63172
ID AAG63172 standard; protein; 285 AA.
XX AC AAG63172;
XX DT 01-OCT-2001 (first entry)
XX DE Amino acid sequence of murine CAR-beta2 and CAR-2 polypeptide.
XX KW Constitutive androstane receptor; CAR; CAR-mediated disorder;
XX KW cholesterol indicator; cholesterol level; hypercholesterolaemia;
XX KW lipid disorder; atherosclerosis; cardiovascular disorder; CAR-2.
XX OS Mus sp.
XX PN WO200151045-A2.
XX PD 19-JUL-2001.
XX PF 12-JAN-2001; 2001WO-US001111.
XX PR 13-JAN-2000; 2000US-0176398P.
XX PA (TULA-) TULARIK INC.
XX PI Lehmann JM, Shiao AK;
XX DR WPI; 2001-476019/51.
XX PT Identifying agent for treating CAR-mediated disorder, involves screening
agent that modulates CAR-mediated intermolecular interaction and
determining if the agent modulates cholesterol level in test mammal.
XX PS Disclosure; Page 72-73; 76pp; English.

CC The specification describes a method for identifying an agent for
treating constitutive androstane receptor (CAR)-mediated disorder/
condition. The method comprises screening candidate agents that modulate
CAR-mediated intermolecular interaction and determining if administration
of the agent to test mammal modulates cholesterol indicator level, or
determining change in cholesterol indicator level in the CAR compromised
mammal to which the agent is administered compared to control. Agents
which modulate CAR-mediated regulation of cholesterol levels, are useful
for treating a CAR-mediated disorder or condition such as
hypercholesterolaemia, lipid disorders, atherosclerosis and
cardiovascular disorders. The agent is useful for testing potential CAR-
mediated disorder treatments, and for studying mechanisms of
hypercholesterolaemia. The present sequence represents a murine CAR-beta2
and CAR-2, which was used in the course of the invention

SQ Sequence 285 AA;
Query Match 51.1%; Score 45; DB 4; Length 285;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QKAAELNOKSKELEQ 17
DB 110 EKALQLNQOQKELVQ 124

RESULT 15
AAP94878
ID AAP94878 standard; protein; 440 AA.
XX AC AAP94878;
XX DT 25-MAR-2003 (revised)
XX DT 02-JUL-1990 (first entry)
XX DE BIK 1 incorporated within the HIS4 gene.
XX KW Yeast promoter; yeast pheromones; FUS-1; BIK-1; HIS-4.
XX OS Saccharomyces cerevisiae.
XX PN WO8810308-A.
XX PD 29-DEC-1988.
XX PF 23-JUN-1988; 88WO-US002129.
XX PR 24-JUN-1987; 87US-00066078.
XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX PI Fink GR, Trueheart J, Elion EA;
XX DR WPI; 1989-023850/03.
XX DR N-PSDB; AAN93100.
XX PT New DNA fragment contg. protein encoding gene and yeast promoter -
controlled by mating pheromone allowing efficient and regulatable
expression.
XX PS Claim 5; Fig 4; 51pp; English.
XX CC ORF lies between the FUS 1 and BIK 1 promoter, BIK 1 running 3' to 5'. At
least one polypeptide may be promoted within a high-copy vector induced
by a-factor for alpha cells, alpha-factor for a-cells using this promoter
system. (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 440 AA;
Query Match 51.1%; Score 45; DB 1; Length 440;
Best Local Similarity 56.2%; Pred. No. 1.8e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 TOKAAELNOKSKELEQ 17
DB 372 TKKLEANEKIKOLEQ 387

RESULT 16
AAR14907
ID AAR14907 standard; protein; 440 AA.
XX AC AAR14907;
XX DT 25-MAR-2003 (revised)
XX DT 03-FEB-1992 (first entry)
XX DE BIK1 protein.
XX KW Pheromone inducible yeast promoter; bilateral karyogamy defect; FUS2;
XX KW FUS1.
XX OS Saccharomyces cerevisiae.
XX PN US5063154-A.
XX PD 05-NOV-1991.
XX PF 24-JUN-1988; 88US-00212270.
XX PR 24-JUN-1987; 87US-00066078.


```
CC sequence of the invention
XX
SQ Sequence 410 AA;

Query Match 50.0%; Score 44; DB 5; Length 410;
Best Local Similarity 38.9%; Pred. No. 2.4e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTQKAAELNQSKELEQ 18
Db 27 PVSQSSDLNQIQKIQK 44

RESULT 24
ADBI0047
ID ADBI0047 standard; protein; 1224 AA.
XX
AC ADBI0047;
XX
DT 20-NOV-2003 (first entry)
XX
DE Alloiococcus otitis antigenic protein SEQ ID NO:3576.
XX
KW Alloiococcus otitidis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection.
XX
OS Alloiococcus otitis.
XX
PN WO2003048304-A2.
XX
PD 12-JUN-2003.
XX
PF 25-NOV-2002; 2002WO-US036123.
XX
PR 29-NOV-2001; 2001US-0333777P.
PR 18-NOV-2002; 2002US-0426742P.
XX
PA (AMHP ) WYETH HOLDINGS CORP.
XX
PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX
DR WPI; 2003-505284/47.
XX
DR N-PSDB; ADBI0050.
XX
PT New Alloiococcus otitidis polynucleotides and polypeptides, useful for
PT treating and diagnosing diseases, drug screening assays and monitoring of
PT effects during drug clinical trials.
XX
PS Claim 33; SEQ ID NO 3576; 1019pp; English.
XX
CC The present invention describes an isolated polynucleotide (I) of
CC Alloiococcus otitidis genomic DNA, which encodes an antigenic protein.
CC Alloiococcus otitidis is a Gram-positive bacterium. Also described: (1)
CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
CC expression vector comprising the novel isolated polynucleotide (I); its
CC complement, degenerate variant or fragment; (3) a genetically engineered
CC host cell, transfected, transformed or infected with the vector of (2);
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
CC composition comprising the polypeptide, its complement, biological
CC equivalent or fragment, or the polynucleotide that is comprised in the
CC expression vector; (6) a pharmaceutical composition comprising the
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
CC of the polypeptides of (1), their biological equivalent or fragment; (8)
CC immunising against Alloiococcus otitidis by administering to a host the
CC immunogenic composition; (9) detecting and/or identifying Alloiococcus
CC otitidis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (I) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Alloiococcus
CC otitidis. The present sequence represents an Alloiococcus otitidis
CC antigen protein from the present invention.
XX
SQ Sequence 1224 AA;

Query Match 50.0%; Score 44; DB 6; Length 1224;
Best Local Similarity 66.7%; Pred. No. 7.8e+02;

Qy 1 PTQKAAELNQSKELEQ 18
Db 27 PVSQSSDLNQIQKIQK 44

RESULT 24
ADBI0047
ID ADBI0047 standard; protein; 1224 AA.
XX
AC ADBI0047;
XX
DT 20-NOV-2003 (first entry)
XX
DE Alloiococcus otitis antigenic protein SEQ ID NO:3576.
XX
KW Alloiococcus otitidis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection.
XX
OS Alloiococcus otitis.
XX
PN WO2003048304-A2.
XX
PD 12-JUN-2003.
XX
PF 25-NOV-2002; 2002WO-US036123.
XX
PR 29-NOV-2001; 2001US-0333777P.
PR 18-NOV-2002; 2002US-0426742P.
XX
PA (AMHP ) WYETH HOLDINGS CORP.
XX
PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX
DR WPI; 2003-505284/47.
XX
DR N-PSDB; ADBI0050.
XX
PT New Alloiococcus otitidis polynucleotides and polypeptides, useful for
PT treating and diagnosing diseases, drug screening assays and monitoring of
PT effects during drug clinical trials.
XX
PS Claim 33; SEQ ID NO 3576; 1019pp; English.
XX
CC The present invention describes an isolated polynucleotide (I) of
CC Alloiococcus otitidis genomic DNA, which encodes an antigenic protein.
CC Alloiococcus otitidis is a Gram-positive bacterium. Also described: (1)
CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
CC expression vector comprising the novel isolated polynucleotide (I); its
CC complement, degenerate variant or fragment; (3) a genetically engineered
CC host cell, transfected, transformed or infected with the vector of (2);
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
CC composition comprising the polypeptide, its complement, biological
CC equivalent or fragment, or the polynucleotide that is comprised in the
CC expression vector; (6) a pharmaceutical composition comprising the
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
CC of the polypeptides of (1), their biological equivalent or fragment; (8)
CC immunising against Alloiococcus otitidis by administering to a host the
CC immunogenic composition; (9) detecting and/or identifying Alloiococcus
CC otitidis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (I) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Alloiococcus
CC otitidis. The present sequence represents an Alloiococcus otitidis
CC antigen protein from the present invention.
XX
SQ Sequence 1224 AA;

Query Match 50.0%; Score 44; DB 5; Length 410;
Best Local Similarity 38.9%; Pred. No. 2.4e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTQKAAELNQSKELEQ 18
Db 27 PVSQSSDLNQIQKIQK 44

RESULT 23
AAOI7656
ID AAOI7656 standard; protein; 410 AA.
XX
AC AAOI7656;
XX
DT 05-AUG-2002 (first entry)
XX
DE H influenzae BASB201 #1.
XX
KW BASB201; otitis media; pneumonia; sinusitis; nosocomial infection;
KW auditive nerve damage; delayed speech learning; vaccine; antibacterial;
KW auditory; antiinflammatory.
XX
OS Haemophilus influenzae.
XX
PH Key Location/Qualifiers
XX Domain 32..300
XX FT /label= KSO_rich_domain
XX FT 301..400
XX FT /label= peptidase_M37-like_domain
XX
PN WO200230967-A2.
XX
PD 18-APR-2002.
XX
PF 05-OCT-2001; 2001WO-EP011561.
XX
PR 13-OCT-2000; 2000GB-00025169.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Thonnard J;
XX
DR WPI; 2002-426267/45.
XX
DR N-PSDB; AAL46624.
XX
PT New isolated non-typeable Haemophilus influenzae BASB201 polypeptides,
PT useful as components of vaccines for treating bacterial infection such as
PT otitis media, delayed speech learning and inflammation of middle ear.
XX
PS Claim 3; Page 87; 90pp; English.
XX
CC The present invention provides the protein and coding sequences of
CC several versions of the BASB201 protein from non-typeable Haemophilus
CC influenzae. These can be used in the production of vaccines against H.
CC influenzae infection, which can cause otitis media in infants and
CC children, pneumonia in elders, sinusitis, nosocomial infections, or
CC invasive diseases, chronic otitis media with hearing loss, fluid
CC accumulation in the middle ear, auditive nerve damage, delayed speech
CC learning, infections of the upper respiratory tract and inflammation of
CC the middle ear. The present sequence is a version of the BASB201 protein
CC sequence of the invention
XX
SQ Sequence 410 AA;

Query Match 50.0%; Score 44; DB 5; Length 410;
Best Local Similarity 38.9%; Pred. No. 2.4e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
```

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 6 AELNOKSKELEQ 17
|:|:|:|:|:|:|
Db 466 ADLNEKEKELQ 477

RESULT 25
AAR21561
ID AAR21561 standard; peptide; 94 AA.
XX AC AAR21561;
XX DT 25-MAR-2003 (revised)
XX DT 04-JUN-1992 (first entry)
XX DE dhaG carboxy terminus.
XX KW Macromolecular system (WMS) operon; initiation of DNA, RNA and;
KW protein synthesis; primase; DNA replication.
XX OS Listeria monocytogenes.
XX PN EP472434-A.
XX PD 26-FEB-1992.
XX XX 22-AUG-1991; 91EP-00307750.
XX XX 23-AUG-1990; 90US-00572191.
XX PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX PA (ABBO) ABBOTT LAB.
XX PI Lupski JR, Katz L;
XX WPI; 1992-066623/09.
XX DR N-PSDB; AAQ21651.
XX XX Anti-sense oligo:nucleotide(s) for treating and diagnosing bacterial in -
PT are complementary to the macromolecular synthesis operon of bacteria they
PT are used to treat.
XX PS Disclosure; Page 12; 48pp; English.
XX CC The dhaG C-terminal protein sequence was deduced from the WMS operon DNA
CC sequence. The dhaG gene product, primase, has several activities
CC including a protein-protein interaction with the primosome complex, a
CC protein-nucleic acid interaction for recognition of the origin, an RNA
CC polymerase activity to synthesise the primer RNA and plays a role in the
CC partitioning of chromosomes. dhaG is the DNA replication initiation gene.
CC See also AAR23002,9 and AAR21562,3. (Updated on 25-MAR-2003 to correct PA
CC field.)
XX SQ Sequence 94 AA;

Query Match 48.9%; Score 43; DB 2; Length 94;
Best Local Similarity 60.0%; Pred. No. 67;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 4 KAELNOKSKELEQ 18
|:|:|:|:|:|:|
Db 74 KRFLKQKKKELEQ 88

RESULT 26
AAG30644
ID AAG30644 standard; protein; 152 AA.
XX AC AAG30644;
XX DT 17-OCT-2000 (first entry)
XX

DE XX Arabidopsis thaliana protein fragment SEQ ID NO: 36672.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX EF1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134321P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139452P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 25-JUN-1999; 99US-0140695P.
PR 26-JUN-1999; 99US-0140823P.
PR 27-JUN-1999; 99US-0140991P.
PR 28-JUN-1999; 99US-0141287P.
PR 29-JUN-1999; 99US-0141842P.
PR 30-JUN-1999; 99US-0142154P.
PR 01-JUL-1999; 99US-0142055P.
PR 02-JUL-1999; 99US-0142390P.
PR 03-JUL-1999; 99US-0142803P.
PR 04-JUL-1999; 99US-0142920P.
PR 05-JUL-1999; 99US-0142977P.
PR 06-JUL-1999; 99US-0143524P.
PR 07-JUL-1999; 99US-0143624P.
PR 08-JUL-1999; 99US-0144005P.
PR 09-JUL-1999; 99US-0144085P.
PR 10-JUL-1999; 99US-0144335P.
PR 11-JUL-1999; 99US-0144335P.
PR 12-JUL-1999; 99US-0144335P.
PR 13-JUL-1999; 99US-0144335P.
PR 14-JUL-1999; 99US-0144335P.
PR 15-JUL-1999; 99US-0144335P.
PR 16-JUL-1999; 99US-0144335P.
PR 17-JUL-1999; 99US-0144335P.
PR 18-JUL-1999; 99US-0144335P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144335P.
PR 21-JUL-1999; 99US-0144335P.
PR 22-JUL-1999; 99US-0144335P.
PR 23-JUL-1999; 99US-0144335P.
PR 24-JUL-1999; 99US-0144335P.
PR 25-JUL-1999; 99US-0144335P.
PR 26-JUL-1999; 99US-0144335P.
PR 27-JUL-1999; 99US-0144335P.
PR 28-JUL-1999; 99US-0144335P.
PR 29-JUL-1999; 99US-0144335P.
PR 30-JUL-1999; 99US-0144335P.
PR 01-AUG-1999; 99US-0144335P.
PR 02-AUG-1999; 99US-0144335P.
PR 03-AUG-1999; 99US-0144335P.
PR 04-AUG-1999; 99US-0144335P.
PR 05-AUG-1999; 99US-0144335P.
PR 06-AUG-1999; 99US-0144335P.
PR 07-AUG-1999; 99US-0144335P.
PR 08-AUG-1999; 99US-0144335P.
PR 09-AUG-1999; 99US-0144335P.
PR 10-AUG-1999; 99US-0144335P.
PR 11-AUG-1999; 99US-0144335P.
PR 12-AUG-1999; 99US-0144335P.
PR 13-AUG-1999; 99US-0144335P.
PR 14-AUG-1999; 99US-0144335P.
PR 15-AUG-1999; 99US-0144335P.
PR 16-AUG-1999; 99US-0144335P.
PR 17-AUG-1999; 99US-0144335P.
PR 18-AUG-1999; 99US-0144335P.
PR 19-AUG-1999; 99US-0144335P.
PR 20-AUG-1999; 99US-0144335P.
PR 21-AUG-1999; 99US-0144335P.
PR 22-AUG-1999; 99US-0144335P.
PR 23-AUG-1999; 99US-0144335P.
PR 24-AUG-1999; 99US-0144335P.
PR 25-AUG-1999; 99US-0144335P.
PR 26-AUG-1999; 99US-0144335P.
PR 27-AUG-1999; 99US-0144335P.
PR 28-AUG-1999; 99US-0144335P.
PR 29-AUG-1999; 99US-0144335P.
PR 30-AUG-1999; 99US-0144335P.

PR 30-AUG-1999; 99US-01511303P.
PR 31-AUG-1999; 99US-01511303P.
PR 01-SEP-1999; 99US-01511303P.
PR 02-SEP-1999; 99US-01511303P.
PR 03-SEP-1999; 99US-01511303P.
PR 04-SEP-1999; 99US-01511303P.
PR 05-SEP-1999; 99US-01511303P.
PR 06-SEP-1999; 99US-01511303P.
PR 07-SEP-1999; 99US-01511303P.
PR 08-SEP-1999; 99US-01511303P.
PR 09-SEP-1999; 99US-01511303P.
PR 10-SEP-1999; 99US-01511303P.
PR 11-SEP-1999; 99US-01511303P.
PR 12-SEP-1999; 99US-01511303P.
PR 13-SEP-1999; 99US-01511303P.
PR 14-SEP-1999; 99US-01511303P.
PR 15-SEP-1999; 99US-01511303P.
PR 16-SEP-1999; 99US-01511303P.
PR 17-SEP-1999; 99US-01511303P.
PR 18-SEP-1999; 99US-01511303P.
PR 19-SEP-1999; 99US-01511303P.
PR 20-SEP-1999; 99US-01511303P.
PR 21-SEP-1999; 99US-01511303P.
PR 22-SEP-1999; 99US-01511303P.
PR 23-SEP-1999; 99US-01511303P.
PR 24-SEP-1999; 99US-01511303P.
PR 25-SEP-1999; 99US-01511303P.
PR 26-SEP-1999; 99US-01511303P.
PR 27-SEP-1999; 99US-01511303P.
PR 28-SEP-1999; 99US-01511303P.
PR 29-SEP-1999; 99US-01511303P.
PR 30-SEP-1999; 99US-01511303P.
PR 01-OCT-1999; 99US-01511303P.
PR 02-OCT-1999; 99US-01511303P.
PR 03-OCT-1999; 99US-01511303P.
PR 04-OCT-1999; 99US-01511303P.
PR 05-OCT-1999; 99US-01511303P.
PR 06-OCT-1999; 99US-01511303P.
PR 07-OCT-1999; 99US-01511303P.
PR 08-OCT-1999; 99US-01511303P.
PR 09-OCT-1999; 99US-01511303P.
PR 10-OCT-1999; 99US-01511303P.
PR 11-OCT-1999; 99US-01511303P.
PR 12-OCT-1999; 99US-01511303P.
PR 13-OCT-1999; 99US-01511303P.
PR 14-OCT-1999; 99US-01511303P.
PR 15-OCT-1999; 99US-01511303P.
PR 16-OCT-1999; 99US-01511303P.
PR 17-OCT-1999; 99US-01511303P.
PR 18-OCT-1999; 99US-01511303P.
PR 19-OCT-1999; 99US-01511303P.
PR 20-OCT-1999; 99US-01511303P.
PR 21-OCT-1999; 99US-01511303P.
PR 22-OCT-1999; 99US-01511303P.
PR 23-OCT-1999; 99US-01511303P.
PR 24-OCT-1999; 99US-01511303P.
PR 25-OCT-1999; 99US-01511303P.
PR 26-OCT-1999; 99US-01511303P.
PR 27-OCT-1999; 99US-01511303P.
PR 28-OCT-1999; 99US-01511303P.
PR 29-OCT-1999; 99US-01511303P.

Query Match 48.9%; Score 43; DB 3; Length 152;
Best Local Similarity 47.1%; Pred. No. 1.1e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 PTKAAELNOKSKELEQ 17
Db 7 PNLAVEIHEEKEQEQ 23

RESULT 27

ABB47539
ID ABB47539 standard; protein; 156 AA.
XX
AC ABB47539;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #243.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX


```
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142300P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142970P.
PR 12-JUL-1999; 99US-0143542P.
PR 13-JUL-1999; 99US-0143624P.
PR 14-JUL-1999; 99US-0144005P.
PR 15-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145154P.
PR 23-JUL-1999; 99US-0145221P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159864P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 48.9%; Score 43; DB 3; Length 219;
Best Local Similarity 47.1%; Pred. No. 1.7e+02;
Matches 8; Conservative

Cy 1 PTKAAELNOKSKELEQ 17
Db 74 PNQLAVEIHEEKEQEQ 90

RESULT 29
AAG30642
ID AAG30642 standard; protein; 228 AA.
XX AC AAG30642;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 36670.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
```

PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128223P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144083P.
PR 16-JUL-1999; 99US-0144086P.
PR 18-JUL-1999; 99US-0144325P.
PR 18-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147915P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149920P.
PR 23-AUG-1999; 99US-0149920P.
PR 23-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154019P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.

```
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161931P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 48.9%; Score 43; DB 3; Length 228;
Best Local Similarity 47.1%; Pred. No. 1.8e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 PTKAAELNQSKELEQ 17
Db 83 PNQLAVEIHEEKEQEQ 99

RESULT 30
AAG41490
ID AAG41490 standard; protein; 323 AA.
XX AC AAG41490;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 51627.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PP 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0132456P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-01444332P.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 17:00:13 ; Search time 143 Seconds
(without alignments)
45.280 Million cell updates

Title: US-10-031-289-1331

Perfect score:

Sequence: 1 PTQAAELNOKSKELEQQ 18

Scoring table: BLOSUM62

```

scoring cable: BSC05062
Gapop 10.0 : Gapext 0.5

```

Searched: 1599051 segs. 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Minimum	DB seq	length:
Maximum	DB seq	length: 200000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published Applications AA:*

- ```

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| Result No. | §     |             |        | SUMMARIES |                      |  | Description          |
|------------|-------|-------------|--------|-----------|----------------------|--|----------------------|
|            | Score | Query Match | Length | DB        | ID                   |  |                      |
| 1          | 88    | 100.0       | 1574   | 16        | US-10-695-499-179    |  | Sequence 179, Appl   |
| 2          | 88    | 100.0       | 1978   | 16        | US-10-695-499-60     |  | Sequence 60, Appl    |
| 3          | 88    | 100.0       | 1981   | 9         | US-09-928-457-38     |  | Sequence 38, Appl    |
| 4          | 88    | 100.0       | 2015   | 14        | US-10-066-551-1      |  | Sequence 1, Appl     |
| 5          | 88    | 100.0       | 2015   | 15        | US-10-282-132A-5772  |  | Sequence 65772, Appl |
| 6          | 88    | 100.0       | 2015   | 17        | US-10-665-980A-1     |  | Sequence 1, Appl     |
| 7          | 67    | 76.1        | 1532   | 16        | US-10-695-499-62     |  | Sequence 62, Appl    |
| 8          | 48    | 54.5        | 241    | 9         | US-09-987-107-24     |  | Sequence 24, Appl    |
| 9          | 45    | 51.1        | 440    | 14        | US-10-369-493-1512   |  | Sequence 1512, A     |
| 10         | 45    | 51.1        | 476    | 16        | US-10-437-963-105696 |  | Sequence 105696,     |
| 11         | 45    | 51.1        | 1030   | 16        | US-10-437-963-183422 |  | Sequence 183422,     |
| 12         | 45    | 51.1        | 1162   | 15        | US-10-424-599-276367 |  | Sequence 276367,     |
| 13         | 45    | 51.1        | 1479   | 16        | US-10-437-963-106132 |  | Sequence 106132,     |

Sequence 181253,  
Sequence 182471,  
Sequence 184860,  
Sequence 182397,  
Sequence 182958,  
Sequence 183619,  
Sequence 184895,  
Sequence 183468,  
Sequence 183049,  
Sequence 181248,  
Sequence 183379,  
Sequence 181403,  
Sequence 181232,  
Sequence 181520,

US-10-437-963-181253  
US-10-437-963-182471  
US-10-437-963-184860  
US-10-437-963-182397  
US-10-437-963-182958  
US-10-437-963-183619  
US-10-437-963-184895  
US-10-437-963-183468  
US-10-437-963-183049  
US-10-437-963-181248  
US-10-437-963-183379  
US-10-437-963-181403  
US-10-437-963-181232  
US-10-437-963-181520

87 42 47.7 815 16  
88 42 47.7 825 16  
89 42 47.7 831 16  
90 42 47.7 833 16  
91 42 47.7 858 16  
92 42 47.7 862 16  
93 42 47.7 866 16  
94 42 47.7 867 16  
95 42 47.7 868 16  
96 42 47.7 870 16  
97 42 47.7 875 16  
98 42 47.7 882 16  
99 42 47.7 884 16  
100 42 47.7 887 16

## ALIGNMENTS

RESULT 1  
US-10-695-499-179  
; Sequence 179, Application US/10695499  
; Publication No. US20040126391A1  
; GENERAL INFORMATION:  
; APPLICANT: Scalato, Enzo  
; APPLICANT: Masignani, Vega  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizza, Mariagrazia  
; APPLICANT: Grandi, Guido  
; TITLE OF INVENTION: Meningococcal Antigens  
; FILE REFERENCE: CHIR0159  
; CURRENT APPLICATION NUMBER: US/10/695,499  
; CURRENT FILING DATE: 2003-10-28  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: PCT/IB99/00103  
; PRIOR FILING DATE: 1999-01-14  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 179  
; LENGTH: 1574  
; TYPE: PRT  
; FEATURE:  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: Description of Artificial Sequence: ORF114-1  
US-10-695-499-179

Query Match 100.0%; Score 88; DB 16; Length 1574;  
Best Local Similarity 100.0%; Pred. No. 0.00021;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18  
Db 959 PTOKAAELNOKSKELEQQ 976

RESULT 2  
US-10-695-499-60  
; Sequence 60, Application US/10695499  
; Publication No. US20040126391A1  
; GENERAL INFORMATION:  
; APPLICANT: Scalato, Enzo  
; APPLICANT: Masignani, Vega  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizza, Mariagrazia  
; APPLICANT: Grandi, Guido  
; TITLE OF INVENTION: Meningococcal Antigens  
; FILE REFERENCE: CHIR0159  
; CURRENT APPLICATION NUMBER: US/10/695,499  
; CURRENT FILING DATE: 2003-10-28  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: PCT/IB99/00103

; PRIOR FILING DATE: 1999-01-14  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 60  
; LENGTH: 1978  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
US-10-695-499-60

Query Match 100.0%; Score 88; DB 16; Length 1978;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18  
Db 959 PTOKAAELNOKSKELEQQ 976

RESULT 3  
US-09-928-457-38  
; Sequence 38, Application US/09928457  
; Patent No. US20020164603A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: DNA, specific proteins and peptides  
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method  
; TITLE OF INVENTION: for obtaining them and their biological application.  
; NUMBER OF SEQUENCES: 99  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (OEB)  
; CURRENT APPLICATION DATA: US/09/928,457  
; FILING DATE: 2001-08-14  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/214,759  
; FILING DATE: 199-12-10  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1981 acids amin,s  
; TYPE: acide amin,  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..1981  
US-09-928-457-38

Query Match 100.0%; Score 88; DB 9; Length 1981;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18  
Db 959 PTOKAAELNOKSKELEQQ 976

RESULT 4  
US-10-066-551-1  
; Sequence 1, Application US/10066551  
; Publication No. US20030100071A1  
; GENERAL INFORMATION:  
; APPLICANT: Apicella, M. A.  
; APPLICANT: Edwards, J. L.  
; APPLICANT: Gibson, B. W.  
; APPLICANT: Scheffler, K.  
; APPLICANT: Brown, E.  
; TITLE OF INVENTION: Vaccine and compositions for the  
; TITLE OF INVENTION: prevention and treatment of Neisserial infections  
; FILE REFERENCE: 875.045US1



; CURRENT APPLICATION NUMBER: US/10/066,551  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: US 60/344,452  
; PRIOR FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: US 60/310,356  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/266,070  
; PRIOR FILING DATE: 2001-01-31  
; NUMBER OF SEQ ID NOS: 12  
; SEQ ID NO 1  
; LENGTH: 2015  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-066-551-1

Query Match 100.0%; Score 88; DB 14; Length 2015;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18  
| | | | | | | | | | | | | | | | | | | |  
Db 959 PTOKAAELNOKSKELEQQ 976

## RESULT 5

US-10-282-122A-65772  
; Sequence 65772, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 65772  
; LENGTH: 2015  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
US-10-282-122A-65772

Query Match 100.0%; Score 88; DB 15; Length 2015;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18  
| | | | | | | | | | | | | | | | | | | |  
Db 959 PTOKAAELNOKSKELEQQ 976

## RESULT 6

US-10-665-990A-1  
; Sequence 1, Application US/10665990A  
; Publication No. US2004025322A1  
; GENERAL INFORMATION:  
; APPLICANT: Apicella, Michael A.  
; APPLICANT: Edwards, Jennifer L.  
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neisseria meningitidis  
; FILE REFERENCE: 17023-031001  
; CURRENT APPLICATION NUMBER: US/10/665,990A  
; CURRENT FILING DATE: 2003-09-19  
; PRIOR APPLICATION NUMBER: US 10/621,184  
; PRIOR FILING DATE: 2003-07-15  
; PRIOR APPLICATION NUMBER: US 10/066,551  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: US 60/344,452  
; PRIOR FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: US 60/310,356  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/266,070  
; PRIOR FILING DATE: 2001-01-31  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2015  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-665-990A-1

Query Match 100.0%; Score 88; DB 17; Length 2015;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18  
| | | | | | | | | | | | | | | | | | | |  
Db 959 PTOKAAELNOKSKELEQQ 976

## RESULT 7

US-10-695-499-62  
; Sequence 62, Application US/10695499  
; Publication No. US20040126391A1  
; GENERAL INFORMATION:  
; APPLICANT: Scalato, Enzo  
; APPLICANT: Masignani, Vega  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizza, Mariagrazia  
; APPLICANT: Grandi, Guido  
; TITLE OF INVENTION: Meningococcal Antigens  
; FILE REFERENCE: CHIR0159  
; CURRENT APPLICATION NUMBER: US/10/695,499  
; CURRENT FILING DATE: 2003-10-28  
; PRIOR APPLICATION NUMBER: US/09/302,626  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: PCT/IB99/00103  
; PRIOR FILING DATE: 1999-01-14  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 62  
; LENGTH: 1532  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
US-10-695-499-62

```

FEATURE:
NAME/KEY: SITE
LOCATION: (68)..(73)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (85)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (296)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (343)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (352)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (377)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (417)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (477)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (665)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (693)..(694)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (781)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (805)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (903)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (922)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (959)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (963)..(965)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (985)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (1157)
OTHER INFORMATION: unknown

```

```

NAME/KEY: SITE
LOCATION: (1164)..(1165)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (1175)..(1176)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (1220)..(1221)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (1249)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (1270)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (1300)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (1447)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (1508)
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: SITE
LOCATION: (1526)
OTHER INFORMATION: unknown
US-10-695-499-62

```

Query Match 76.1%; Score 67; DB 16; Length 1532;  
 Best Local Similarity 82.4%; Pred. No. 0.32;  
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy 2 TQAAELNQSKELEQQ 18
 |||||
Db 960 TQXXXLNQSKELEQQ 976

```

```

RESULT 8
US-09-987-107-24
Sequence 24, Application US/09987107
Patent No. US20020156007A1
GENERAL INFORMATION:
APPLICANT: GRAVERSEN, Jonas
APPLICANT: MOESTRUP, Soren
TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
FILE REFERENCE: GRAVERSEN1A
CURRENT APPLICATION NUMBER: US/09/987,107
CURRENT FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/264,022
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: DK PA2001 00057
PRIOR FILING DATE: 2001-01-15
PRIOR APPLICATION NUMBER: DK PA2000 01682
PRIOR FILING DATE: 2000-11-10
NUMBER OF SEQ ID NOS: 91
SOFTWARE: Patent in version 3.1
SEQ ID NO 24
LENGTH: 241
TYPE: PRT
ORGANISM: Erinaceus europaeus
US-09-987-107-24

```

Query Match 54.5%; Score 48; DB 9; Length 241;  
 Best Local Similarity 68.8%; Pred. No. 32;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 QKAAELNQKSKELEQ 18  
||| ||| : |||  
Db 131 QKAQELQKAGELGQ 146

## RESULT 9

US-10-369-493-1512  
; Sequence 1512, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 1512  
; LENGTH: 440  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-10-369-493-1512

Query Match 51.1%; Score 45; DB 14; Length 440;  
Best Local Similarity 56.2%; Pred. No. 1.7e+02;  
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TQKAAELNQKSKELEQ 17  
||| ||| : |||  
Db 372 TKLEEAENEKIQLEQ 387

## RESULT 10

US-10-437-963-105696  
; Sequence 105696, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 105696  
; LENGTH: 476  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_102915C.1.pep  
US-10-437-963-105696

Query Match 51.1%; Score 45; DB 16; Length 476;  
Best Local Similarity 41.2%; Pred. No. 1.9e+02;  
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PTQAAELNQKSKELEQ 17  
||| ||| : |||

Db 431 PTAKAAEMRQARAIKE 447

## RESULT 11

US-10-437-963-183422  
; Sequence 183422, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 183422  
; LENGTH: 1030  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_80515C.1.pep  
US-10-437-963-183422

Query Match 51.1%; Score 45; DB 16; Length 1030;  
Best Local Similarity 64.3%; Pred. No. 4.4e+02;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNQKSKELE 16  
||| ||| : |||  
Db 816 ORAAELEARAKELE 829

## RESULT 12

US-10-424-599-276367  
; Sequence 276367, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 276367  
; LENGTH: 1182  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1182)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_91579C.1.pep  
US-10-424-599-276367

Query Match 51.1%; Score 45; DB 15; Length 1182;  
Best Local Similarity 47.1%; Pred. No. 5.2e+02;  
Matches 8; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TQKAAELNQKSKELEQ 18  
||| ||| : |||  
Db 183 SRKSEFDRKSQLESEQ 199

RESULT 17

US-10-425-115-354199  
; Sequence 354199, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 354199  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_86200C.1.pep  
US-10-425-115-354199

Query Match 50.0%; Score 44; DB 17; Length 225;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PTOKAAELNQSKSELEQ 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 82 PKTKAAKRNERRKKEKQQ 99

## RESULT 18

US-10-425-115-354200  
; Sequence 354200, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 354200  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_86201C.1.pep  
US-10-425-115-354200

Query Match 50.0%; Score 44; DB 17; Length 225;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PTOKAAELNQSKSELEQ 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 82 PKTKAAKRNERRKKEKQQ 99

## RESULT 19

US-10-425-114-57118  
; Sequence 57118, Application US/10425114  
; Publication No. US20040034889A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 57118  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700244127\_FLI.pep  
US-10-425-114-57118

Query Match 50.0%; Score 44; DB 15; Length 250;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PTOKAAELNQSKSELEQ 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 106 PKTKAAKRNERRKKEKQQ 123

## RESULT 20

US-10-369-493-2962  
; Sequence 2962, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2962  
; LENGTH: 304  
; TYPE: PRT  
; ORGANISM: Thermotoga maritima  
US-10-369-493-2962

Query Match 50.0%; Score 44; DB 14; Length 304;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 QKAAELNQSKSELEQ 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 91 QKIDELISKELEQ 106

## RESULT 21

US-10-399-091-2  
; Sequence 2, Application US/10399091  
; Publication No. US20040047875A1  
; GENERAL INFORMATION:  
; APPLICANT: Thonnard, Joelle  
; TITLE OF INVENTION: No. US20040047875A1e1 Compounds  
; FILE REFERENCE: BM45420  
; CURRENT APPLICATION NUMBER: US/10/399,091  
; CURRENT FILING DATE: 2003-04-11  
; PRIOR APPLICATION NUMBER: PCT/EP01/11561  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: GB0025169.4  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 2
; LENGTH: 410
; TYPE: PRT
; ORGANISM: non-typeable Haemophilus influenzae
US-10-399-091-2

Query Match 50.0%; Score 44; DB 15; Length 410;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Oy 1 PTOKAAELNOKSKELEQQ 18
 |:::|:::|:::|
Db 27 PVSQSSDLNQLQKQIKQQ 44

RESULT 22
US-10-399-091-4
; Sequence 6, Application US/10399091
; Publication No. US20040047875A1
; GENERAL INFORMATION:
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: No. US20040047875A1el Compounds
; FILE REFERENCE: BM45420
; CURRENT APPLICATION NUMBER: US/10/399,091
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: PCT/EP01/11561
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: GB0025169.4
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 410
; TYPE: PRT
; ORGANISM: non-typeable Haemophilus influenzae
US-10-399-091-4

Query Match 50.0%; Score 44; DB 15; Length 410;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Oy 1 PTOKAAELNOKSKELEQQ 18
 |:::|:::|:::|
Db 27 PVSQSSDLNQLQKQIKQQ 44

RESULT 23
US-10-399-091-6
; Sequence 6, Application US/10399091
; Publication No. US20040047875A1
; GENERAL INFORMATION:
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: No. US20040047875A1el Compounds
; FILE REFERENCE: BM45420
; CURRENT APPLICATION NUMBER: US/10/399,091
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: PCT/EP01/11561
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: GB0025169.4
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 410
; TYPE: PRT
; ORGANISM: non-typeable Haemophilus influenzae
US-10-399-091-6

Query Match 50.0%; Score 44; DB 15; Length 410;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Oy 1 PTOKAAELNOKSKELEQQ 18
```

```
Db 27 PVSQSSDLNQLQKQIKQQ 44
 |:::|:::|:::|

RESULT 24
US-10-399-091-8
; Sequence 8, Application US/10399091
; Publication No. US20040047875A1
; GENERAL INFORMATION:
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: No. US20040047875A1el Compounds
; FILE REFERENCE: BM45420
; CURRENT APPLICATION NUMBER: US/10/399,091
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: PCT/EP01/11561
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: GB0025169.4
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 410
; TYPE: PRT
; ORGANISM: non-typeable Haemophilus influenzae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 372
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-399-091-8

Query Match 50.0%; Score 44; DB 15; Length 410;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Oy 1 PTOKAAELNOKSKELEQQ 18
 |:::|:::|:::|
Db 27 PVSQSSDLNQLQKQIKQQ 44

RESULT 25
US-10-369-493-6955
; Sequence 6955, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6955
; LENGTH: 926
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6955

Query Match 50.0%; Score 44; DB 14; Length 926;
Best Local Similarity 38.9%; Pred. No. 5.6e+02;
Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Oy 1 PTOKAAELNOKSKELEQQ 18
 |:::|:::|:::|
Db 824 PVKASEIAEKARDLERE 841

RESULT 26
```

US-10-767-701-44430  
; Sequence 44430, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 44430  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(200)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C45320\_1.pep  
US-10-767-701-44430

Query Match 48.9%; Score 43; DB 16; Length 117;  
Best Local Similarity 62.5%; Pred. No. 82;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 QKAAELNOKSKELRQQ 18  
|||:|||||:  
Db 54 EKLAENQKLKELQK 69

RESULT 27  
US-10-437-963-195907  
; Sequence 195907, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 195907  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_9180C.1.pep  
US-10-437-963-195907

Query Match 48.9%; Score 43; DB 16; Length 163;  
Best Local Similarity 57.1%; Pred. No. 1.2e+02;  
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 KAAELNOKSKELQ 17  
|||:|||||:  
Db 93 KMSLNTRSKELDE 106

RESULT 28  
US-10-424-599-169323  
; Sequence 169323, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 169323  
; LENGTH: 200  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(200)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_123913C.1.pep  
US-10-424-599-169323

Query Match 48.9%; Score 43; DB 15; Length 200;  
Best Local Similarity 58.8%; Pred. No. 1.5e+02;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 TOKAAELNOKSKELQQ 18  
|||:|||||:  
Db 102 TGKAAINQISPESEQE 118

RESULT 29  
US-10-739-930-6093  
; Sequence 6093, Application US/10739930  
; Publication No. US20040216190A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-21(53377)B  
; CURRENT APPLICATION NUMBER: US/10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 6093  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C22035\_1.p  
US-10-739-930-6093

Query Match 48.9%; Score 43; DB 17; Length 228;  
Best Local Similarity 47.1%; Pred. No. 1.7e+02;  
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 PTKAAELNOKSKELQ 17  
|||:|||||:  
Db 83 PNQLAVEIHEEKEQEQ 99

RESULT 30  
US-10-739-930-5722  
; Sequence 5722, Application US/10739930  
; Publication No. US20040216190A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-21(53377)B  
; CURRENT APPLICATION NUMBER: US/10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 5722  
; LENGTH: 323  
; TYPE: PRT

```
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C126388_1.p
US-10-739-930-5722
```

```
Query Match 48.9%; Score 43; DB 17; Length 323;
Best Local Similarity 60.0%; Pred. NO. 2.5e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 4 KAEINOKSKELEQ 18
 |: ||||| :||
Db 30 KSTEINQKIGDLEQ 44
```

```
Search completed: December 30, 2004, 17:12:24
Job time : 145 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:51:42 ; Search time 37 Seconds  
(without alignments)  
32.263 Million cell updates/sec

Title: US-10-031-289-1331

Perfect score: 88

Sequence: 1 PTKAAELNQSKELEQQ 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 88    | 100.0       | 1574   | 4  | US-09-302-626B-179   |
| 2          | 88    | 100.0       | 1978   | 4  | US-09-302-626B-60    |
| 3          | 67    | 76.1        | 1532   | 4  | US-09-302-626B-62    |
| 4          | 46    | 52.3        | 810    | 3  | US-09-540-824-25     |
| 5          | 44    | 50.0        | 43     | 3  | US-09-015-030-6      |
| 6          | 43    | 48.9        | 312    | 4  | US-09-248-796A-16239 |
| 7          | 43    | 48.9        | 326    | 4  | US-09-248-796A-16951 |
| 8          | 43    | 48.9        | 385    | 4  | US-09-270-767-44065  |
| 9          | 42    | 47.7        | 108    | 4  | US-09-513-999C-5914  |
| 10         | 42    | 47.7        | 329    | 2  | US-09-004-502-3      |
| 11         | 42    | 47.7        | 329    | 3  | US-09-360-125-3      |
| 12         | 42    | 47.7        | 359    | 4  | US-09-328-352-7097   |
| 13         | 42    | 47.7        | 414    | 5  | PCT-US93-03077-3     |
| 14         | 42    | 47.7        | 532    | 4  | US-09-352-391A-27654 |
| 15         | 42    | 47.7        | 1093   | 5  | PCT-US93-03077-1     |
| 16         | 42    | 47.7        | 2472   | 4  | US-09-538-092-1312   |
| 17         | 41    | 46.6        | 89     | 2  | US-08-690-011A-52    |
| 18         | 41    | 46.6        | 89     | 3  | US-09-299-495F-64    |
| 19         | 41    | 46.6        | 194    | 4  | US-09-134-000C-4448  |
| 20         | 41    | 46.6        | 315    | 4  | US-09-328-352-4290   |
| 21         | 41    | 46.6        | 468    | 4  | US-09-489-039A-7971  |
| 22         | 41    | 46.6        | 915    | 4  | US-09-540-236-3026   |
| 23         | 41    | 46.6        | 1282   | 4  | US-09-543-681A-5419  |
| 24         | 40    | 45.5        | 82     | 4  | US-09-513-999C-5870  |
| 25         | 40    | 45.5        | 148    | 4  | US-09-270-767-58267  |
| 26         | 40    | 45.5        | 180    | 4  | US-09-149-476-401    |
| 27         | 40    | 45.5        | 347    | 2  | US-09-004-502-1      |

|     |      |      |      |   |                      |                   |
|-----|------|------|------|---|----------------------|-------------------|
| 28  | 40   | 45.5 | 347  | 3 | US-09-360-125-1      | Sequence 1, Appl  |
| 29  | 40   | 45.5 | 475  | 4 | US-09-248-796A-14568 | Sequence 14568, A |
| 30  | 40   | 45.5 | 694  | 3 | US-08-559-397A-31    | Sequence 31, Appl |
| 31  | 40   | 45.5 | 1024 | 4 | US-09-562-737-86     | Sequence 86, Appl |
| 32  | 40   | 45.5 | 1132 | 4 | US-09-248-796A-15026 | Sequence 15026, A |
| 33  | 40   | 45.5 | 1485 | 4 | US-09-543-681A-5367  | Sequence 5367, Ap |
| 34  | 39.5 | 44.9 | 46   | 4 | US-09-898-554-29     | Sequence 29, Appl |
| 35  | 39.5 | 44.9 | 46   | 4 | US-09-898-554-32     | Sequence 32, Appl |
| 36  | 39.5 | 44.9 | 46   | 4 | US-09-898-554-34     | Sequence 34, Appl |
| 37  | 39.5 | 44.9 | 165  | 4 | US-09-898-554-24     | Sequence 24, Appl |
| 38  | 39.5 | 44.9 | 207  | 4 | US-09-898-554-26     | Sequence 26, Appl |
| 39  | 39.5 | 44.9 | 284  | 4 | US-09-914-259-46     | Sequence 46, Appl |
| 40  | 39.5 | 44.9 | 363  | 4 | US-09-898-554-20     | Sequence 20, Appl |
| 41  | 39.5 | 44.9 | 1388 | 3 | US-09-572-191-2      | Sequence 2, Appl  |
| 42  | 39.5 | 44.9 | 1388 | 3 | US-09-723-262-2      | Sequence 2, Appl  |
| 43  | 39.5 | 44.9 | 1388 | 3 | US-09-723-219-2      | Sequence 2, Appl  |
| 44  | 39   | 44.3 | 27   | 4 | US-09-544-664B-26    | Sequence 26, Appl |
| 45  | 39   | 44.3 | 34   | 2 | US-08-690-011A-53    | Sequence 53, Appl |
| 46  | 39   | 44.3 | 34   | 2 | US-08-690-011A-57    | Sequence 57, Appl |
| 47  | 39   | 44.3 | 34   | 3 | US-09-299-495F-53    | Sequence 53, Appl |
| 48  | 39   | 44.3 | 34   | 3 | US-09-299-495F-57    | Sequence 57, Appl |
| 49  | 39   | 44.3 | 82   | 4 | US-09-621-976-5189   | Sequence 5189, Ap |
| 50  | 39   | 44.3 | 98   | 2 | US-08-690-011A-19    | Sequence 19, Appl |
| 51  | 39   | 44.3 | 101  | 4 | US-09-543-681A-4289  | Sequence 4289, Ap |
| 52  | 39   | 44.3 | 179  | 4 | US-09-134-000C-6423  | Sequence 6423, Ap |
| 53  | 39   | 44.3 | 213  | 4 | US-08-529-055-47     | Sequence 47, Appl |
| 54  | 39   | 44.3 | 216  | 4 | US-09-252-991A-27173 | Sequence 27173, A |
| 55  | 39   | 44.3 | 230  | 4 | US-09-270-767-41815  | Sequence 41815, A |
| 56  | 39   | 44.3 | 234  | 4 | US-09-270-767-43273  | Sequence 43273, A |
| 57  | 39   | 44.3 | 256  | 4 | US-09-248-796A-13647 | Sequence 13647, A |
| 58  | 39   | 44.3 | 258  | 4 | US-09-107-532A-4002  | Sequence 4002, Ap |
| 59  | 39   | 44.3 | 330  | 4 | US-09-301-666A-10    | Sequence 10, Appl |
| 60  | 39   | 44.3 | 330  | 4 | US-09-301-217-10     | Sequence 10, Appl |
| 61  | 39   | 44.3 | 352  | 4 | US-09-710-279-216    | Sequence 216, App |
| 62  | 39   | 44.3 | 431  | 4 | US-09-134-000C-5352  | Sequence 5352, Ap |
| 63  | 39   | 44.3 | 438  | 4 | US-09-248-796A-18215 | Sequence 18215, A |
| 64  | 39   | 44.3 | 450  | 4 | US-09-248-796A-20424 | Sequence 20424, A |
| 65  | 39   | 44.3 | 503  | 4 | US-09-634-238-267    | Sequence 267, App |
| 66  | 39   | 44.3 | 553  | 4 | US-09-922-364A-3     | Sequence 3, Appl  |
| 67  | 39   | 44.3 | 557  | 4 | US-09-254-590-3      | Sequence 3, Appl  |
| 68  | 39   | 44.3 | 557  | 4 | US-09-922-364A-20    | Sequence 20, Appl |
| 69  | 39   | 44.3 | 557  | 4 | US-09-254-590-20     | Sequence 20, Appl |
| 70  | 39   | 44.3 | 646  | 4 | US-09-710-279-676    | Sequence 676, App |
| 71  | 39   | 44.3 | 731  | 3 | US-09-115-446-2      | Sequence 2, Appl  |
| 72  | 39   | 44.3 | 731  | 4 | US-09-275-252A-10    | Sequence 10, Appl |
| 73  | 39   | 44.3 | 731  | 4 | US-09-565-590-2      | Sequence 2, Appl  |
| 74  | 39   | 44.3 | 732  | 4 | US-09-922-364A-43    | Sequence 43, Appl |
| 75  | 39   | 44.3 | 732  | 4 | US-09-254-590-43     | Sequence 43, Appl |
| 76  | 39   | 44.3 | 736  | 4 | US-09-922-364A-47    | Sequence 47, Appl |
| 77  | 39   | 44.3 | 736  | 4 | US-09-254-590-47     | Sequence 47, Appl |
| 78  | 39   | 44.3 | 737  | 1 | US-08-188-582-16     | Sequence 16, Appl |
| 79  | 39   | 44.3 | 737  | 1 | US-08-646-715-16     | Sequence 16, Appl |
| 80  | 39   | 44.3 | 966  | 4 | US-09-688-188B-154   | Sequence 154, App |
| 81  | 39   | 44.3 | 966  | 4 | US-09-291-417D-155   | Sequence 155, App |
| 82  | 39   | 44.3 | 968  | 4 | US-09-688-188B-107   | Sequence 107, App |
| 83  | 39   | 44.3 | 968  | 4 | US-09-688-188B-155   | Sequence 155, App |
| 84  | 39   | 44.3 | 968  | 4 | US-09-291-417D-107   | Sequence 107, App |
| 85  | 39   | 44.3 | 968  | 4 | US-09-291-417D-155   | Sequence 155, App |
| 86  | 39   | 44.3 | 1275 | 4 | US-08-426-630-49     | Sequence 49, Appl |
| 87  | 39   | 44.3 | 1341 | 4 | US-09-248-796A-13074 | Sequence 13074, A |
| 88  | 39   | 44.3 | 8991 | 4 | US-08-714-741-32     | Sequence 32, Appl |
| 89  | 38.5 | 43.8 | 133  | 4 | US-09-248-796A-27920 | Sequence 27920, A |
| 90  | 38.5 | 43.8 | 176  | 4 | US-09-710-279-362    | Sequence 362, App |
| 91  | 38   | 43.2 | 26   | 2 | US-08-690-011A-3     | Sequence 3, Appl  |
| 92  | 38   | 43.2 | 26   | 2 | US-09-299-495F-3     | Sequence 3, Appl  |
| 93  | 38   | 43.2 | 33   | 2 | US-08-690-011A-38    | Sequence 38, Appl |
| 94  | 38   | 43.2 | 33   | 2 | US-09-299-495F-38    | Sequence 38, Appl |
| 95  | 38   | 43.2 | 34   | 3 | US-08-690-011A-54    | Sequence 54, Appl |
| 96  | 38   | 43.2 | 34   | 3 | US-09-299-495F-54    | Sequence 54, Appl |
| 97  | 38   | 43.2 | 84   | 3 | US-09-299-495F-11    | Sequence 11, Appl |
| 98  | 38   | 43.2 | 85   | 4 | US-09-134-000C-4168  | Sequence 4168, Ap |
| 99  | 38   | 43.2 | 86   | 3 | US-09-299-495F-17    | Sequence 17, Appl |
| 100 | 38   | 43.2 | 86   | 3 | US-09-299-495F-52    | Sequence 52, Appl |

## ALIGNMENTS

## RESULT 1

US-09-302-626B-179  
; Sequence 179, Application US/09302626B  
; Patent No. 6709660  
; GENERAL INFORMATION:  
; APPLICANT: Scalato, Enzo  
; APPLICANT: Masignani, Vega  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizza, Mariagrazia  
; TITLE OF INVENTION: Meningococcal Antigens  
; FILE REFERENCE: CHIR0159  
; CURRENT APPLICATION NUMBER: US/09/302,626B  
; CURRENT FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: PCT/IB99/00103  
; PRIOR FILING DATE: 1999-01-14  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 179  
; LENGTH: 1574  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: ORF114-1  
US-09-302-626B-179

Query Match 100.0%; Score 88; DB 4; Length 1574;  
Best Local Similarity 100.0%; Pred. No. 3.3e-05;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTQAAELNOKSKELEQ 18  
Db 959 PTQAAELNOKSKELEQ 976

## RESULT 2

US-09-302-626B-60  
; Sequence 60, Application US/09302626B  
; Patent No. 6709660  
; GENERAL INFORMATION:  
; APPLICANT: Scalato, Enzo  
; APPLICANT: Masignani, Vega  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizza, Mariagrazia  
; APPLICANT: Grandi, Guido  
; TITLE OF INVENTION: Meningococcal Antigens  
; FILE REFERENCE: CHIR0159  
; CURRENT APPLICATION NUMBER: US/09/302,626B  
; CURRENT FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: PCT/IB99/00103  
; PRIOR FILING DATE: 1999-01-14  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 60  
; LENGTH: 1978  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
US-09-302-626B-60

Query Match 100.0%; Score 88; DB 4; Length 1978;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTQAAELNOKSKELEQ 18  
Db 959 PTQAAELNOKSKELEQ 976

## RESULT 3

US-09-302-626B-62  
; Sequence 62, Application US/09302626B  
; Patent No. 6709660  
; GENERAL INFORMATION:  
; APPLICANT: Scalato, Enzo  
; APPLICANT: Masignani, Vega  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizza, Mariagrazia  
; APPLICANT: Grandi, Guido  
; TITLE OF INVENTION: Meningococcal Antigens  
; FILE REFERENCE: CHIR0159  
; CURRENT APPLICATION NUMBER: US/09/302,626B  
; CURRENT FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: PCT/IB99/00103  
; PRIOR FILING DATE: 1999-01-14  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 62  
; LENGTH: 1532  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (68)..(73)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (85)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (296)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (343)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (352)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (377)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (417)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (665)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (693)..(694)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (781)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (805)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (903)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (922)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (959)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (963)..(965)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (985)

; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1157)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1164)..(1165)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1175)..(1176)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1220)..(1221)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1249)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1270)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1300)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1447)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1508)  
; OTHER INFORMATION: unknown  
; NAME/KEY: SITE  
; LOCATION: (1526)  
; OTHER INFORMATION: unknown  
; US-302-626B-62

Query Match 76.1%; Score 67; DB 4; Length 1532;  
Best Local Similarity 82.4%; Pred. No. 0.055;  
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TQAAELNQSKLEEQ 18  
Db 960 TQXXXLNQSKLEEQ 976

RESULT 4  
US-09-540-824-25  
; Sequence 25, Application US/09540824  
; Patent No. 6383753  
; GENERAL INFORMATION:  
; APPLICANT: Thiele, Dennis  
; TITLE OF INVENTION: No. 6383753el Yeast and Mammalian Regulators of Cell Proliferation  
; FILE REFERENCE: UM-04266  
; CURRENT APPLICATION NUMBER: US/09/540,824  
; CURRENT FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 25  
; LENGTH: 810  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
; US-09-540-824-25

Query Match 52.3%; Score 46; DB 3; Length 810;  
Best Local Similarity 56.2%; Pred. No. 48;  
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 QKAAELNQSKLEEQ 18  
Db 294 EKAAEEKRELEEQ 309

RESULT 5  
US-09-015-030-6  
; Sequence 6, Application US/09015030

; Patent No. 6303317  
; GENERAL INFORMATION:  
; APPLICANT: Alber, Thomas C.  
; APPLICANT: Sharma, Victoria A.  
; APPLICANT: Nautiyal, Shivani  
; TITLE OF INVENTION: Peptide Probes and Methods for Making the Same  
; FILE REFERENCE: A-64988-1/RPT/TAL  
; CURRENT APPLICATION NUMBER: US/09/015,030  
; CURRENT FILING DATE: 1998-01-28  
; PRIOR APPLICATION NUMBER: 60/036,219  
; PRIOR FILING DATE: 1997-01-28  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: yeast  
; US-09-015-030-6

Query Match 50.0%; Score 44; DB 3; Length 43;  
Best Local Similarity 50.0%; Pred. No. 4.2;  
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 QKAAELNQSKLEEQ 18  
Db 14 EKYELKKNKLEQE 29

RESULT 6  
US-09-248-796A-16239  
; Sequence 16239, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 16239  
; LENGTH: 312  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; US-09-248-796A-16239

Query Match 48.9%; Score 43; DB 4; Length 312;  
Best Local Similarity 72.7%; Pred. No. 50;  
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 7 ELNQSKLEEQ 17  
Db 99 ELNKKLEQ 109

RESULT 7  
US-09-248-796A-16951  
; Sequence 16951, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409

```
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16951
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16951

Query Match 48.9%; Score 43; DB 4; Length 326;
Best Local Similarity 53.3%; Pred. No. 52;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 4 KAAELNOKSKELEQ 18
Db 216 KQELDARSKKIEQ 230

RESULT 8
US-09-270-767-44065
; Sequence 44065, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1998-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44065
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44065

Query Match 48.9%; Score 43; DB 4; Length 385;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PTKAAELNOKSKELEQ 18
Db 46 PAKSPSLDKSKKYDRQ 63

RESULT 9
US-09-513-999C-5914
; Sequence 5914, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duciart, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2 REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5914
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 82
; OTHER INFORMATION: Xaa=Lys or Arg
US-09-513-999C-5914

Query Match 47.7%; Score 42; DB 4; Length 108;
Best Local Similarity 57.1%; Pred. No. 23;

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELE 16
Db 94 RKAELERKERELQ 107

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELE 16
Db 94 RKAELERKERELQ 107

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

RESULT 10
US-09-004-502-3
; Sequence 3, Application US/09004502
; Patent No. 5962263
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN MEMBRANE RECYCLING
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,502
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0456 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 329 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT01
; CLONE: 412453
US-09-004-502-3

Query Match 47.7%; Score 42; DB 2; Length 329;
Best Local Similarity 57.1%; Pred. No. 75;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELE 16
Db 94 RKAELERKERELQ 107

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

RESULT 11
US-09-360-125-3
; Sequence 3, Application US/09360125
; Patent No. 6235715
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN MEMBRANE RECYCLING
```

;; TITLE OF INVENTION: PROTEINS  
;; NUMBER OF SEQUENCES: 5  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
;; STREET: 3174 Porter Dr.  
;; CITY: Palo Alto  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94304  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/360,125  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/004,502  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Billings, Lucy J.  
;; REGISTRATION NUMBER: 36,749  
;; REFERENCE/DOCKET NUMBER: PP-0456 US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 650-855-0555  
;; TELEFAX: 650-845-4166  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 329 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; LIBRARY: BRSTN0701  
;; CLONE: 412453  
;; US-09-360-125-3

Query Match 47.7%; Score 42; DB 3; Length 329;  
Best Local Similarity 57.1%; Pred. No. 75;  
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKLE 16  
Db 94 RKAELERKERELQ 107

RESULT 12  
US-09-328-352-7097  
;; Sequence 7097, Application US/09328352  
;; Patent No. 6562958  
;; GENERAL INFORMATION:  
;; APPLICANT: Gary L. Breton et al.  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
;; FILE REFERENCE: GTC99-03PA  
;; CURRENT APPLICATION NUMBER: US/09/328,352  
;; CURRENT FILING DATE: 1999-06-04  
;; NUMBER OF SEQ ID NOS: 8252  
;; SEQ ID NO 7097  
;; LENGTH: 359  
;; TYPE: PRT  
;; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7097

Query Match 47.7%; Score 42; DB 4; Length 359;  
Best Local Similarity 57.1%; Pred. No. 82;  
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 5 AAELNOKSKELEQ 18  
Db 278 ANELNQRLEQ 291

RESULT 13  
PCT-US93-03077-3  
;; Sequence 3, Application PC/TUS9303077  
;; GENERAL INFORMATION:  
;; APPLICANT: Board of Regents, The University of Texas System  
;; APPLICANT: Gaynor, Richard B.  
;; APPLICANT: Wu, Foon Kin  
;; TITLE OF INVENTION: PROTEIN CELLULAR FACTOR USEFUL FOR  
;; TITLE OF INVENTION: REGULATING GENE EXPRESSION  
;; NUMBER OF SEQUENCES: 7  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Arnold White & Durkee  
;; STREET: P.O. Box 4433  
;; CITY: Houston  
;; STATE: Texas  
;; COUNTRY: USA  
;; ZIP: 77210  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US93/03077  
;; FILING DATE: 19930331  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/862,025  
;; FILING DATE: April 2, 1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kammerer, Patricia A.  
;; REGISTRATION NUMBER: 29,775  
;; REFERENCE/DOCKET NUMBER: UTFD270PCT  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 713-787-1540  
;; TELEFAX: 713-749-2679  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 414 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
;; PCT-US93-03077-3

Query Match 47.7%; Score 42; DB 5; Length 414;  
Best Local Similarity 61.5%; Pred. No. 96;  
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AELNOKSKELEQ 18  
Db 155 AKLNKKVKELEE 167

RESULT 14  
US-09-252-991A-27654  
;; Sequence 27654, Application US/09252991A  
;; Patent No. 6551795  
;; GENERAL INFORMATION:  
;; APPLICANT: Marc J. Rubenfield et al.  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
;; FILE REFERENCE: 107196.136  
;; CURRENT APPLICATION NUMBER: US/09/252,991A  
;; CURRENT FILING DATE: 1999-02-18  
;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27654  
; LENGTH: 592  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27654

Query Match 47.7%; Score 42; DB 4; Length 592;  
Best Local Similarity 46.7%; Pred. No. 1.4e+02;  
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 4 KAAELNQKSKELEQ 18  
Db 220 RAAELQKQKQLEQ 234  
:|||||:|:|:|

RESULT 15  
PCT-US93-03077-1  
; Sequence 1, Application PC/TUS9303077  
; GENERAL INFORMATION:  
; APPLICANT: Board of Regents, The University of Texas System  
; APPLICANT: Gaynor, Richard B.  
; TITLE OF INVENTION: PROTEIN CELLULAR FACTOR USEFUL FOR  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/03077  
; FILING DATE: 19930331  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/862,025  
; FILING DATE: April 2, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kammerer, Patricia A.  
; REGISTRATION NUMBER: 29,775  
; REFERENCE/DOCKET NUMBER: UTFFD270PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713-787-1540  
; TELEFAX: 713-749-2679  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1093 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
PCT-US93-03077-1

Query Match 47.7%; Score 42; DB 5; Length 1093;  
Best Local Similarity 61.5%; Pred. No. 2.7e+02;  
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AELNQKSKELEQ 18  
Db 591 AKLNKVKLEEE 603  
:|||||:|:|:|

RESULT 16  
US-09-538-092-1312  
; Sequence 1312, Application US/09538092

; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CurapatSeqFormatter Version 0.9  
; SEQ ID NO 1312  
; LENGTH: 2472  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number Q13813  
US-09-538-092-1312

Query Match 47.7%; Score 42; DB 4; Length 2472;  
Best Local Similarity 56.2%; Pred. No. 6.5e+02;  
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 QKAAELNQKSKELEQ 18  
Db 1645 QKSAEKSQKLKANKQ 1660  
:|||||:|:|:|

RESULT 17  
US-08-690-011A-52  
; Sequence 52, Application US/08690011A  
; Patent No. 5942433  
; GENERAL INFORMATION:  
; APPLICANT: VINSON, Charles R.  
; APPLICANT: KRYLOV, Dmitry  
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN  
; TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A  
; TITLE OF INVENTION: CELLULAR PROTEIN  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/690,011A  
; FILING DATE: 31-JUL-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/001,654  
; FILING DATE: 31-JUL-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/018,496  
; FILING DATE: 29-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Serunian, Leslie A.  
; REGISTRATION NUMBER: 35,353  
; REFERENCE/DOCKET NUMBER: 2026-4199US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)758-4800  
; TELEFAX: (212)751-6849  
; INFORMATION FOR SEQ ID NO: 52:



```
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7971
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7971

Query Match 46.6%; Score 41; DB 4; Length 468;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 KAAELNOKSKELEQ 18
Db 318 KNPDLVKPKLEQR 332

RESULT 22
US-09-540-236-3026
; Sequence 3026, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3026
; LENGTH: 915
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3026

Query Match 46.6%; Score 41; DB 4; Length 915;
Best Local Similarity 43.8%; Pred. No. 3.2e+02;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELEQ 18
Db 765 QRVAQLSDKRRHLEQ 780

RESULT 23
US-09-543-681A-5419
; Sequence 5419, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5419
; LENGTH: 1282
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5419

Query Match 46.6%; Score 41; DB 4; Length 1282;
Best Local Similarity 41.2%; Pred. No. 4.6e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQ 17
Db 484 PSQRAELRRREALQQ 500

RESULT 24
```

```
US-09-513-999C-5870
; Sequence 5870, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5870
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5870

Query Match 45.5%; Score 40; DB 4; Length 82;
Best Local Similarity 43.8%; Pred. No. 34;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELEQ 18
Db 60 KQFELNKAELDRL 75

RESULT 25
US-09-270-767-58267
; Sequence 58267, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58267
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-58267

Query Match 45.5%; Score 40; DB 4; Length 148;
Best Local Similarity 60.0%; Pred. No. 65;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKEL 15
Db 30 PTQKPYSPNOKKPL 44

RESULT 26
US-09-149-476-401
; Sequence 401, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002p1
; CURRENT APPLICATION NUMBER: US/09/149.476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
```



[illegible]

;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,593  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,614  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/043,578  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,576  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/047,501  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/043,670  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/056,632  
;; EARLIER FILING DATE: 1997-08-22  
;; EARLIER APPLICATION NUMBER: 60/056,664  
;; EARLIER FILING DATE: 1997-08-22  
;; EARLIER APPLICATION NUMBER: 60/056,876  
;; EARLIER FILING DATE: 1997-08-22  
;; EARLIER APPLICATION NUMBER: 60/056,881  
;; EARLIER FILING DATE: 1997-08-22  
;; EARLIER APPLICATION NUMBER: 60/056,909  
;; EARLIER FILING DATE: 1997-08-22  
;; EARLIER APPLICATION NUMBER: 60/056,908  
;; EARLIER FILING DATE: 1997-08-22  
;; EARLIER APPLICATION NUMBER: 60/048,964  
;; EARLIER FILING DATE: 1997-06-06  
;; EARLIER APPLICATION NUMBER: 60/057,650  
;; EARLIER FILING DATE: 1997-09-05  
;; EARLIER APPLICATION NUMBER: 60/056,884  
;; EARLIER FILING DATE: 1997-08-22  
;; EARLIER APPLICATION NUMBER: 60/057,669  
;; EARLIER FILING DATE: 1997-09-05  
;; EARLIER APPLICATION NUMBER: 60/049,610  
;; EARLIER FILING DATE: 1997-06-13  
;; EARLIER APPLICATION NUMBER: 60/061,060  
;; EARLIER FILING DATE: 1997-10-02

Query Match 45.5%; Score 40; DB 4; Length 180;  
Best Local Similarity 43.8%; Pred. No. 80;  
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELEQQ 18  
Db 101 KQELNRKAELEDR 116

RESULT 27  
US-09-004-502-1  
; Sequence 1, Application US/09004502  
; Patent No. 5962263  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Shah, Purvi  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN MEMBRANE RECYCLING  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette

;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/004,502  
;; FILING DATE: Filed Herewith  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Billings, Lucy J.  
;; REGISTRATION NUMBER: 36,749  
;; REFERENCE/DOCKET NUMBER: PF-0456 US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 650-855-0555  
;; TELEFAX: 650-845-4166  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 347 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; LIBRARY: TONGTUT01  
;; CLONE: 980615  
US-09-004-502-1

Query Match 45.5%; Score 40; DB 2; Length 347;  
Best Local Similarity 43.8%; Pred. No. 1.6e+02;  
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELEQQ 18  
Db 101 KQELNRKAELEDR 116

RESULT 28  
US-09-360-125-1  
; Sequence 1, Application US/09360125  
; Patent No. 6235715  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Shah, Purvi  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN MEMBRANE RECYCLING  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/360,125  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/004,502  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0456 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555

; TELEFAX: 650-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: TONGTUT01  
; CLONE: 980615  
US-09-360-125-1

Query Match 45.5%; Score 40; DB 3; Length 347;  
Best Local Similarity 43.8%; Pred. No. 1.6e+02;  
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSKELEQ 18  
Db 101 KQBELNKAELDOR 116

## RESULT 29

US-09-248-796A-14568  
; Sequence 14568, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 14568  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-14568

Query Match 45.5%; Score 40; DB 4; Length 475;  
Best Local Similarity 64.3%; Pred. No. 2.3e+02;  
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 KAAELNOKSKELEQ 17  
Db 342 RAKELHAKSVELEQ 355

## RESULT 30

US-08-559-397A-31  
; Sequence 31, Application US/08559397A  
; Patent No. 6083713  
; GENERAL INFORMATION:  
; APPLICANT: Manly, Susan P.  
; APPLICANT: Kozlowski, Michael R.  
; APPLICANT: Neve, Rachael L.  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF  
; TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/559,397A  
; FILING DATE: 15-NOV-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 6013-135  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-559-397A-31

Query Match 45.5%; Score 40; DB 3; Length 694;  
Best Local Similarity 44.4%; Pred. No. 3.4e+02;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQ 18  
Db 354 PTSKPAQARSLSKELNEK 371

Search completed: December 30, 2004, 17:01:30  
Job time : 39 secs

